

NOVEL CARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES
THEREFOR

Related Applications

This application claims priority to U.S. Provisional Application No. 60/111,938, filed on December 11, 1998, U.S. Patent Application No.: 09/291,839, filed on April 14, 1999, and U.S. Patent Application No.: 09/458,457, filed on December 10, 1999, incorporated herein in their entirety by this reference.

Background of the Invention

Phosphate tightly associated with protein has been known since the late nineteenth century. Since then, a variety of covalent linkages of phosphate to proteins have been found. The most common involve esterification of phosphate to serine, threonine, and tyrosine with smaller amounts being linked to lysine, arginine, histidine, aspartic acid, glutamic acid, and cysteine. The occurrence of phosphorylated proteins implies the existence of one or more protein kinases capable of phosphorylating amino acid residues on proteins, and also of protein phosphatases capable of hydrolyzing phosphorylated amino acid residues on proteins.

Protein kinases play critical roles in the regulation of biochemical and morphological changes associated with cellular growth and division (D'Urso, G. *et al.* (1990) *Science* 250: 786-791; Birchmeier, C. *et al.* (1993) *Bioessays* 15: 185-189). They serve as growth factor receptors and signal transducers and have been implicated in cellular transformation and malignancy (Hunter, T. *et al.* (1992) *Cell* 70: 375-387; Posada, J. *et al.* (1992) *Mol. Biol. Cell* 3: 583-592; Hunter, T. *et al.* (1994) *Cell* 79: 573-582). For example, protein kinases have been shown to participate in the transmission of signals from growth-factor receptors (Sturgill, T. W. *et al.* (1988) *Nature* 344: 715-718; Gomez, N. *et al.* (1991) *Nature* 353: 170-173), control of entry of cells into mitosis (Nurse, P. (1990) *Nature* 344: 503-508; Maller, J. L. (1991) *Curr. Opin. Cell Biol.* 3: 269-275) and regulation of actin bundling (Husain-Chishti, A. *et al.* (1988) *Nature* 334: 718-721). Protein kinases can be divided into two main groups based on either amino acid sequence similarity or specificity for either serine/threonine or tyrosine residues. A small number of dual-specificity kinases are structurally like the serine/threonine-specific group. Within the broad classification, kinases can be further sub-divided into families whose members share a higher degree of catalytic domain amino acid sequence identity and also have similar biochemical properties. Most protein kinase family members also share structural features outside the kinase domain that reflect their particular cellular roles. These include regulatory domains that control kinase

activity or interaction with other proteins (Hanks, S.K. *et al.* (1988) *Science* 241: 42-52). For example, kinases which contain ankyrin repeat domains have been identified, such as the Integrin-linked kinase (ILK).

ILK is an ankyrin repeat containing serine-threonine protein kinase which interacts with integrin β_1 and β_3 heterodimeric transmembrane glycoprotein subunit cytoplasmic domains. Integrins communicate with cell surface and cytoplasmic molecules such as cytoskeletal and catalytic signaling proteins (Hannigan G.E. *et al.* (1981) *Nature* 379:91-96, Schwartz M.A. *et al.* (1995) *Annu. Rev. Cell Dev. Biol.* 11:549-599). Overexpression of ILK increases the expression of cyclin A, cyclin D₁ and Cdk4 proteins by overriding the adhesion-dependent regulation of cell cycle progression through G₁ into S phase. This activity suggests that ILK may be an important regulator of integrin-mediated cell cycle progression (Radeva G. *et al.* (1997) *J. Biol. Chem.* 272:13937-13944). Overexpression of ILK also stimulates fibronectin matrix assembly in epithelial cells (Wu C. *et al.* (1998) *J. Biol. Chem.* 273:528-536).

Summary of the Invention

The present invention is based, at least in part, on the discovery of novel ankyrin repeat containing kinases, referred herein as "Cardiac-related Ankyrin-Repeat Protein Kinase" ("CARK") nucleic acid and protein molecules. The CARK molecules of the present invention are useful as modulating agents for regulating a variety of cellular processes, *e.g.*, cardiac cellular processes. Accordingly, in one aspect, this invention provides isolated nucleic acid molecules encoding CARK proteins or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection of CARK-encoding nucleic acids.

In one embodiment, a CARK nucleic acid molecule of the invention is at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or more identical to the nucleotide sequence (*e.g.*, to the entire length of the nucleotide sequence) shown in SEQ ID NO:1 or 3 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a complement thereof.

In one embodiment, a CARK nucleic acid molecule of the invention is at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 82%, 85%, 90%, 95%, 98%, or more identical to the nucleotide sequence (*e.g.*, to the entire length of the nucleotide sequence) shown in SEQ ID NO:7 or 9 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a complement thereof.

In a preferred embodiment, the isolated nucleic acid molecule includes the nucleotide sequence shown SEQ ID NO:1 or 3, or a complement thereof. In another embodiment, the nucleic acid molecule includes SEQ ID NO:3 and nucleotides 1-47 of SEQ ID NO:1. In another embodiment, the nucleic acid molecule includes SEQ ID NO:3 and

nucleotides 2553-3025 of SEQ ID NO:1. In another preferred embodiment, the nucleic acid molecule consists of the nucleotide sequence shown in SEQ ID NO:1 or 3. In another preferred embodiment, the nucleic acid molecule includes a fragment of at least 350, 467, 762, 918, 1236, or 1275 nucleotides (*e.g.*, contiguous nucleotides) of the nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof.

In a preferred embodiment, the isolated nucleic acid molecule includes the nucleotide sequence shown SEQ ID NO:7 or 9, or a complement thereof. In another embodiment, the nucleic acid molecule includes SEQ ID NO:9 and nucleotides 1-60 of SEQ ID NO:7. In another embodiment, the nucleic acid molecule includes SEQ ID NO:9 and nucleotides 2566-3026 of SEQ ID NO:7. In another preferred embodiment, the nucleic acid molecule consists of the nucleotide sequence shown in SEQ ID NO:7 or 9. In another preferred embodiment, the nucleic acid molecule includes a fragment of at least 2962 nucleotides (*e.g.*, 2962 contiguous nucleotides) of the nucleotide sequence of SEQ ID NO:7 or 9, or a complement thereof.

In another embodiment, a CARK nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence sufficiently homologous to the amino acid sequence of SEQ ID NO:2 or 8, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In a preferred embodiment, a CARK nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to the entire length of the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another preferred embodiment, a CARK nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 92%, 95%, 98% or more homologous to the entire length of the amino acid sequence of SEQ ID NO:8, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

In another preferred embodiment, an isolated nucleic acid molecule encodes the amino acid sequence of human CARK. In yet another preferred embodiment, the nucleic acid molecule includes a nucleotide sequence encoding a protein having the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In yet another preferred embodiment, the nucleic acid molecule is at least 350, 467, 762, 918, 1236, or 1275 nucleotides in length. In a further preferred embodiment, the nucleic acid molecule is at least 350, 467, 762, 918, 1236, or 1275 nucleotides in length and encodes a protein having a CARK activity (as described herein).

In another preferred embodiment, an isolated nucleic acid molecule encodes the amino acid sequence of rat CARK. In yet another preferred embodiment, the nucleic acid molecule includes a nucleotide sequence encoding a protein having the amino acid sequence of SEQ ID NO:8, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In yet another preferred embodiment, the nucleic acid molecule is at least 2962 nucleotides in length. In a further preferred embodiment, the nucleic acid molecule is at least 2962 nucleotides in length and encodes a protein having a CARK activity (as described herein).

Another embodiment of the invention features nucleic acid molecules, preferably CARK nucleic acid molecules, which specifically detect CARK nucleic acid molecules relative to nucleic acid molecules encoding non-CARK proteins. For example, in one embodiment, such a nucleic acid molecule is at least 250-300, 300-335, 339, 339-350, 350-400, 400-450, 467, 467-500, 500-550, or 550-600, 600-750, 762, 762-800, 800-900, 918, 918-1000, 1000-1200, 1236, 1275, 1275-1400, 1400-1600, 1600-1800, 1800-2000, 2000-2400, 2400-2800, 2800-2900, 2962, or more nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence shown in SEQ ID NO:1, 3, 7, or 9, or, the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a complement thereof.

In preferred embodiments, the nucleic acid molecules are at least 15 (*e.g.*, contiguous) nucleotides in length and hybridize under stringent conditions to nucleotides 1-47, 94-294, 318-338, 1145-1364, 1833-1921, 2051-2267, 2290-2543, or 3018-3025 of SEQ ID NO:1. In other preferred embodiments, the nucleic acid molecules comprise nucleotides 1-47, 94-294, 318-338, 1145-1364, 1833-1921, 2051-2267, 2290-2543, or 3018-3025 of SEQ ID NO:1. In other preferred embodiments, the nucleic acid molecules consist of nucleotides 1-47, 94-294, 318-338, 1145-1364, 1833-1921, 2051-2267, 2290-2543, or 3018-3025 of SEQ ID NO:1.

In other preferred embodiments, the nucleic acid molecule encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 8, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising SEQ ID NO:1 or 3, or SEQ ID NO:7 or 9, respectively, under stringent conditions.

Another embodiment of the invention provides an isolated nucleic acid molecule which is antisense to a CARK nucleic acid molecule, *e.g.*, the coding strand of a CARK nucleic acid molecule.

Another aspect of the invention provides a vector comprising a CARK nucleic acid molecule. In certain embodiments, the vector is a recombinant expression vector. In another embodiment, the invention provides a host cell containing a vector of the invention.

In yet another embodiment, the invention provides a host cell containing a nucleic acid molecule of the invention. The invention also provides a method for producing a protein, preferably a CARK protein, by culturing in a suitable medium, a host cell, *e.g.*, a mammalian host cell such as a non-human mammalian cell, of the invention containing a recombinant expression vector, such that the protein is produced.

Another aspect of this invention features isolated or recombinant CARK proteins and polypeptides. In one embodiment, the isolated protein, preferably a CARK protein, includes at least one ankyrin repeat domain, and preferably two, three, four, five, six, seven, eight, or, most preferably, nine or more ankyrin repeat domains. In another embodiment, the isolated protein, preferably a CARK protein, includes at least one protein kinase domain. In a preferred embodiment, the protein, preferably a CARK protein, includes at least one ankyrin repeat domain and preferably two, three, four, five, six, seven, eight, or, most preferably, nine or more ankyrin repeat domains, and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 92%, 95%, 98% or more homologous to the amino acid sequence of SEQ ID NO:2 or 8, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another preferred embodiment, the protein, preferably a CARK protein, includes at least one protein kinase domain and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 92%, 95%, 98% or more homologous to the amino acid sequence of SEQ ID NO:2 or 8, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another preferred embodiment, the protein, preferably a CARK protein, includes at least one LXCXE motif and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 92%, 95%, 98% or more homologous to the amino acid sequence of SEQ ID NO:2 or 8, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In yet another preferred embodiment, the protein, preferably a CARK protein, includes at least one ankyrin repeat domain and preferably two, three, four, five, six, seven, eight, or, most preferably, nine or more ankyrin repeat domains, at least one protein kinase domain, and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 92%, 95%, 98% or more homologous to the amino acid sequence of SEQ ID NO:2 or 8, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In a preferred embodiment, a CARK protein includes at least one or more of the following domains and/or motifs: an ankyrin repeat domain, a kinase domain or a LXCXE motif, and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 92%, 95%, 98% or more homologous to the amino acid sequence of SEQ ID NO:2 or 8, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

In another preferred embodiment, the protein, preferably a CARK protein, includes at least one ankyrin repeat domain and preferably two, three, four, five, six, seven, eight, or, most preferably, nine or more ankyrin repeat domains and has a CARK activity, as

described herein. In yet another preferred embodiment, the protein, preferably a CARK

5 protein, includes a protein kinase domain and has a CARK activity, as described herein. In

another preferred embodiment, the protein, preferably a CARK protein, includes a LXCXE motif and has a CARK activity, as described herein. In a further preferred embodiment, the

protein, preferably a CARK protein, includes at least one ankyrin repeat domain and

preferably two, three, four, five, six, seven, eight, or, most preferably, nine or more ankyrin

10 repeat domains, at least one protein kinase domain, and has a CARK activity, as described

herein. In a preferred embodiment, a CARK protein includes at least one or more of the following domains and/or motifs: an ankyrin repeat domain, a kinase domain or a LXCXE motif, and has a CARK activity, as described herein.

In yet another preferred embodiment, the protein, preferably a CARK protein, includes at

15 least one ankyrin repeat domain and preferably two, three, four, five, six, seven, eight, or,

most preferably, nine or more ankyrin repeat domains and is encoded by a nucleic acid

molecule having a nucleotide sequence which hybridizes under stringent hybridization

conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1,

3, 7, or 9. In a further embodiment, the protein, preferably a CARK protein, includes at least

20 one protein kinase domain and is encoded by a nucleic acid molecule having a nucleotide

sequence which hybridizes under stringent hybridization conditions to a nucleic acid

molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 7, or 9. In a further

embodiment, the protein, preferably a CARK protein, includes at least one LXCXE motif

and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes

25 under stringent hybridization conditions to a nucleic acid molecule comprising the

nucleotide sequence of SEQ ID NO:1, 3, 7, or 9. In another preferred embodiment, the

protein, preferably a CARK protein, includes at least one ankyrin repeat domain and

preferably two, three, four, five, six, seven, eight, or, most preferably, nine or more ankyrin

repeat domains, at least one protein kinase domain, and is encoded by a nucleic acid

30 molecule having a nucleotide sequence which hybridizes under stringent hybridization

conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1,

3, 7, or 9. In a preferred embodiment, a CARK protein includes at least one or more of the

following domains and/or motifs: an ankyrin repeat domain, a kinase domain or a LXCXE

motif, and is encoded by a nucleic acid molecule having a nucleotide sequence which

35 hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising

the nucleotide sequence of SEQ ID NO:1, 3, 7, or 9.

In another embodiment, the invention features fragments of the protein having the amino acid sequence of SEQ ID NO:2 or 8, wherein the fragment comprises at least 15

amino acids (*e.g.*, contiguous amino acids) of the amino acid sequence of SEQ ID NO:2 or 8 or an amino acid sequence encoded by the DNA insert of the plasmid deposited with the ATCC as Accession Number _____. In one embodiment, the protein comprises amino acid residues 463 to 716 of SEQ ID NO:2 or 8. In another embodiment, the invention comprises amino acid residues 411-415 of SEQ ID NO:2 or 8. In yet another embodiment, the protein, preferably a CARK protein, has the amino acid sequence of SEQ ID NO:2 or 8.

In another embodiment, the invention features an isolated protein, preferably a CARK protein, which is encoded by a nucleic acid molecule consisting of a nucleotide sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 82%, 85%, 90%, 95%, 98% or more homologous to a nucleotide sequence of SEQ ID NO:1, 3, 7, or 9, or a complement thereof. This invention further features an isolated protein, preferably a CARK protein, which is encoded by a nucleic acid molecule consisting of a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 7, or 9, or a complement thereof.

The proteins of the present invention or portions thereof, *e.g.*, biologically active portions thereof, can be operatively linked to a non-CARK polypeptide (*e.g.*, heterologous amino acid sequences) to form fusion proteins. The invention further features antibodies, such as monoclonal or polyclonal antibodies, that specifically bind proteins of the invention, preferably CARK proteins. In addition, the CARK proteins or biologically active portions thereof can be incorporated into pharmaceutical compositions, which optionally include pharmaceutically acceptable carriers.

In another aspect, the present invention provides a method for detecting the presence of a CARK nucleic acid molecule, protein or polypeptide in a biological sample by contacting the biological sample with an agent capable of detecting a CARK nucleic acid molecule, protein or polypeptide such that the presence of a CARK nucleic acid molecule, protein or polypeptide is detected in the biological sample.

In another aspect, the present invention provides a method for detecting the presence of CARK activity in a biological sample by contacting the biological sample with an agent capable of detecting an indicator of CARK activity such that the presence of CARK activity is detected in the biological sample.

In another aspect, the invention provides a method for modulating CARK activity comprising contacting a cell capable of expressing CARK with an agent that modulates CARK activity such that CARK activity in the cell is modulated. In one embodiment, the agent inhibits CARK activity. In another embodiment, the agent stimulates CARK activity. In one embodiment, the agent is an antibody that specifically binds to a CARK protein. In another embodiment, the agent modulates expression of CARK by modulating transcription of a CARK gene or translation of a CARK mRNA. In yet another embodiment, the agent is

a nucleic acid molecule having a nucleotide sequence that is antisense to the coding strand of a CARK mRNA or a CARK gene.

In one embodiment, the methods of the present invention are used to treat a subject having a disorder characterized by aberrant or unwanted CARK protein or nucleic acid expression or activity by administering an agent which is a CARK modulator to the subject. In one embodiment, the CARK modulator is a CARK protein. In another embodiment the CARK modulator is a CARK nucleic acid molecule. In yet another embodiment, the CARK modulator is a peptide, peptidomimetic, or other small molecule. In a preferred embodiment, the disorder characterized by aberrant or unwanted CARK protein or nucleic acid expression is a cardiovascular disorder.

The present invention also provides a diagnostic assay for identifying the presence or absence of a genetic alteration characterized by at least one of (i) aberrant modification or mutation of a gene encoding a CARK protein; (ii) mis-regulation of the gene; and (iii) aberrant post-translational modification of a CARK protein, wherein a wild-type form of the gene encodes a protein with a CARK activity.

In another aspect the invention provides a method for identifying a compound that binds to or modulates the activity of a CARK protein, by providing an indicator composition comprising a CARK protein having CARK activity, contacting the indicator composition with a test compound, and determining the effect of the test compound on CARK activity in the indicator composition to identify a compound that modulates the activity of a CARK protein.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

Brief Description of the Drawings

Figure 1 depicts the cDNA sequence and predicted amino acid sequence of human CARK. The nucleotide sequence corresponds to nucleic acids 1 to 3025 of SEQ ID NO:1. The amino acid sequence corresponds to amino acids 1 to 835 of SEQ ID NO: 2. The coding region without the 5' and 3' untranslated regions of the human CARK gene is shown in SEQ ID NO:3.

Figure 2 depicts a structural, hydrophobicity, and antigenicity analysis of the human CARK protein.

Figure 3 depicts an alignment of the CARK protein with the A. Thaliana kinase 2 (Accession No.Z97337), A. Thaliana kinase 3 (Accession No. AC003113), Arabidopsis thaliana (Accession No. AL031135), C. elegans kinase (Accession No. AF024491), D. discoideum protein tyrosine kinase (Accession No. A35670), D. discoideum protein tyrosine kinase (Accession No. U01064), H. sapiens serine/threonine kinase (Accession No.

Z48615), human raf1 (Accession No. W13107), human Raf1 kinase (Accession No. R98215), and soybean kinase (Accession No. M67449) using the Clustal method with a PAM250 residue weight table.

Figure 4 depicts the results of a search which was performed against the HMM database in which ankyrin repeat domains and a protein kinase domain were identified in the human CARK protein.

Figure 5 depicts the cDNA sequence and predicted amino acid sequence of rat CARK. The nucleotide sequence corresponds to nucleic acids 1 to 3026 of SEQ ID NO:7. The amino acid sequence corresponds to amino acids 1 to 835 of SEQ ID NO: 8. The coding region without the 5' and 3' untranslated regions of the rat CARK gene is shown in SEQ ID NO:9.

Figure 6 depicts the results of a search which was performed against the HMM database in which ankyrin repeat domains and a protein kinase domain were identified in the rat CARK protein.

Figure 7 depicts a global alignment of the human CARK nucleic acid sequence with the rat CARK nucleic acid sequence using the GAP program in the GCG software package, using a nwsgapdna matrix a gap weight of 12 and a length weight of 4. The results showed a 82.2% identity between the two sequences.

Figure 8 depicts a global alignment of the human CARK protein with the rat CARK protein using the GAP program in the GCG software package, using a Blosum 62 matrix and a gap weight of 12 and a length weight of 4. The results showed a 91.4% identity between the two sequences.

Figure 9 is a graph depicting CARK expression in human congestive heart failure tissue samples as assessed by PCR.

Figure 10 depicts the *in vitro* autophosphorylation of GST-CARK and GST-CARK-KD polypeptides on tyrosine residue(s) as assessed by western blotting using an anti-phosphotyrosine antibody.

Figure 11 depicts the autophosphorylation of myc-CARK polypeptides in 293 cell transfectants.

Figure 12 is a graph depicting the attenuation of serum induced or phenylephrine induced *de novo* protein synthesis in cardiac myocytes expressing a dominant negative CARK polypeptide.

Detailed Description of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as "Cardiac-related Ankyrin-Repeat Protein Kinase" or "CARK" nucleic acid and protein molecules, which play a role in or function in signaling pathways associated with cellular growth or differentiation. In one embodiment, the CARK molecules modulate

the activity of one or more proteins involved in cellular growth or differentiation, *e.g.*, cardiac cell growth or differentiation. In another embodiment, the CARK molecules of the present invention are capable of modulating the phosphorylation state of a CARK molecule or one or more proteins involved in cellular growth or differentiation, *e.g.*, cardiac cell growth or differentiation. In a preferred embodiment, the CARK molecules of the invention are capable of autophosphorylation, *e.g.*, autophosphorylation on a tyrosine residue (*e.g.*, amino acid residue 771 of SEQ ID NO: 2 or 8).

In a preferred embodiment, the CARK molecules are protein kinases, *e.g.*, serine/threonine protein kinases, which are expressed and/or function in cells of the cardiovascular system, *e.g.*, cells of the heart, the blood vessels, and/or the blood. In another preferred embodiment, the CARK molecules are dual specificity protein kinases.

As used herein, the term "protein kinase" includes a protein or polypeptide which is capable of modulating its own phosphorylation state or the phosphorylation state of another protein or polypeptide. Protein kinases (described in, for example, Hanks S.K. *et al.* (1995) *FASEB J.* 9:576-596) can have a specificity for (*i.e.*, a specificity to phosphorylate) serine/threonine residues, tyrosine residues, or both serine/threonine and tyrosine residues, *e.g.*, the dual specificity kinases. There are a number of conserved regions in the catalytic domain of protein kinases. One of these regions, located in the N-terminal extremity of the catalytic domain, is a glycine-rich stretch of residues in the vicinity of a lysine residue, which is believed to be involved in ATP binding. Another region, located in the central part of the catalytic domain, contains a conserved aspartic acid residue which is important for the catalytic activity of the enzyme (Knighton D.R. *et al.* (1991) *Science* 253:407-414). Two signature patterns have been described for this region: one specific for serine/ threonine kinases and one for tyrosine kinases. Protein kinase polypeptides of the present invention preferably include one of the following consensus sequences:

[LIV] -G- {P} -G- {P} - [FYWMGSTNH] - [SGA] - {PW} - [LIVCAT] - {PD} -x
[GSTACLIVMFY] -x (5, 18) - [LIVMFYWCSTAR] - [AIVP] - [LIVMFAGCKR] -K
(SEQ ID NO:4)
[K binds ATP]

[LIVMFYC] -x- [HY] -x-D- [LIVMFY] -K-x(2) -N- [LIVMFYCT] (3)
(SEQ ID NO:5)
[D is an active site residue]

[LIVMFYC] -x- [HY] -x-D- [LIVMFY] - [RSTAC] -x(2) -N- [LIVMFYC] (3)

(SEQ ID NO:6)

[D is an active site residue]

- 5 The signature patterns or consensus patterns described herein are described according to the following designation: all amino acids are indicated according to their universal single letter designations; "x" designates any amino acid; x (n) designates n number of amino acids, *e.g.*, x(2) designates any 2 amino acids, *e.g.*, x(1-3) designates any of one to three amino acids; and, amino acids in brackets indicates any one of the amino acids within the brackets, *e.g.*,
10 [HY] indicates any one of either H (histidine) or Y (tyrosine).

Protein kinases play a role in signaling pathways associated with cellular growth and differentiation. For example, protein kinases are involved in the regulation of signal transmission from cellular receptors, *e.g.*, growth-factor receptors; entry of cells into mitosis; and the regulation of cytoskeleton function, *e.g.*, actin bundling. Thus, the CARK
15 molecules of the present invention may be involved in: 1) the regulation of transmission of signals from cellular receptors, *e.g.*, cardiac cell growth factor receptors; 2) the modulation of the entry of cells, *e.g.*, cardiac precursor cells, into mitosis; 3) the modulation of cellular differentiation; 4) the modulation of cell death; and 5) the regulation of cytoskeleton function, *e.g.*, actin bundling.

20 Accordingly, the CARK molecules of the present invention may be involved in cellular signal transduction pathways that modulate cardiac cell activity. As used herein, a "cardiac cell activity" or "cardiac cell function" includes cell proliferation, differentiation, migration, hypertrophy and expression of genes associated with cardiac cell function (*e.g.*, ANF, contractile proteins (*e.g.*, myosin, actin), ion channels), as well as cellular processes
25 that contribute to myogenesis, cardiogenesis, and the physiological role of cardiac cells (*e.g.*, contraction and generation of force to pump blood). Thus, the CARK molecules, by participating in cellular signal transduction pathways, may modulate cell behavior and act as targets and therapeutic agents for controlling cardiac cell proliferation, differentiation, hypertrophy and migration.

30 Inhibition or over stimulation of the activity of protein kinases involved in signaling pathways associated with cellular growth can lead to perturbed cellular growth, which can in turn lead to cellular growth related disorders. As used herein, a "cellular growth related disorder" includes a disorder, disease, or condition characterized by a deregulation, *e.g.*, an upregulation or a downregulation, of cellular growth. Cellular growth deregulation may be
35 due to a deregulation of cellular proliferation, cell cycle progression, cellular differentiation and/or cellular hypertrophy. Examples of cellular growth related disorders include cardiovascular disorders such as heart failure, congenital heart disease, cardiac myocyte hypertrophy, hypertension, atrial fibrillation, dilated cardiomyopathy, idiopathic

cardiomyopathy, or angina; and proliferative disorders and/or differentiative disorders such as cancer, *e.g.*, melanoma, prostate cancer, cervical cancer, breast cancer, colon cancer, leukemia, carcinoma, or sarcoma.

As used herein, the term "cardiovascular disorder" includes a disease, disorder, or state involving the cardiovascular system, *e.g.*, the heart, the blood vessels, and/or the blood. A cardiovascular disorder can be caused by an imbalance in arterial pressure, a malfunction of the heart, or an occlusion of a blood vessel, *e.g.*, by a thrombus. Examples of such disorders include congenital heart defects (*e.g.*, atrioventricular canal defects), hypertension, atherosclerosis, coronary artery spasm, coronary artery disease, valvular disease, ischemia reperfusion injury, restenosis, arterial inflammation, vascular wall remodeling, ventricular remodeling, rapid ventricular pacing, coronary microembolism, tachycardia, bradycardia, pressure overload, aortic bending, coronary artery ligation, vascular heart disease, long-QT syndrome, congestive heart failure, sinus node dysfunction, atrial flutter, myocardial infarction, coronary artery spasm, arrhythmias, and cardiomyopathies.

As used herein, the term "congestive heart failure" includes a condition characterized by a diminished capacity of the heart to supply the oxygen demands of the body. Symptoms and signs of congestive heart failure include diminished blood flow to the various tissues of the body, accumulation of excess blood in the various organs, *e.g.*, when the heart is unable to pump out the blood returned to it by the great veins, exertional dyspnea, fatigue, and/or peripheral edema, *e.g.*, peripheral edema resulting from left ventricular dysfunction. Congestive heart failure may be acute or chronic. The manifestation of congestive heart failure usually occurs secondary to a variety of cardiac or systemic disorders that share a temporal or permanent loss of cardiac function. Examples of such disorders include hypertension, coronary artery disease, valvular disease, and cardiomyopathies, *e.g.*, hypertrophic, dilative, or restrictive cardiomyopathies. Congestive heart failure is described in, for example, Cohn J.N. *et al.* (1998) *American Family Physician* 57:1901-04, the contents of which are incorporated herein by reference.

CARK-associated or related disorders include disorders associated with the atrioventricular canal defects (AVCD) locus on human chromosome 1, *e.g.*, congenital heart defects (Sheffield, VC *et al.* (1997) *Human Molecular Genetics*, 6:117-121). CARK-associated or related disorders also include disorders of tissues in which CARK is expressed, *e.g.*, heart and skeletal muscle.

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as CARK protein and nucleic acid molecules, which comprise a family of molecules having certain conserved structural and functional features.

The term "family" when referring to the protein and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence

homology as defined herein. Such family members can be naturally or non-naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin, as well as other, distinct proteins of human origin or alternatively, can contain homologues of non-human origin. Members of a family may also have common functional characteristics.

For example, the family of CARK proteins comprise at least one, and preferably two, three, four, five, six, seven, eight, or most preferably, nine or more ankyrin repeat domains. As used herein, the term "ankyrin repeat domain" includes a protein domain involved in protein:protein interactions, having an amino acid sequence of about 20-40 amino acid residues and having a bit score for the alignment of the sequence to the ankyrin repeat domain (HMM) of at least 1. Preferably, an ankyrin repeat domain includes at least about 25-40, more preferably about 25-35 amino acid residues, or most preferably about 30-35 amino acids and has a bit score for the alignment of the sequence to the ankyrin repeat domain (HMM) of at least 3, 5, 10, 20, 30, 40, 50, or greater. The ankyrin repeat domain (HMM) has been assigned the PFAM Accession PF00023 (<http://genome.wustl.edu/Pfam/.html>). Ankyrin repeats are described in, for example, Otto E. *et al.* (1991) *J. Biol. Chem.* 114:241-253, Hatada E.N. *et al.* (1992) *PNAS USA* 89:2489-2493, and Blank V.P. *et al.* (1992) *Trends Genet.* 8:144-149, the contents of which are incorporated herein by reference.

To identify the presence of an ankyrin repeat domain in a CARK protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of HMMs (*e.g.*, the Pfam database, release 2.1) using the default parameters (http://www.sanger.ac.uk/Software/Pfam/HMM_search). For example, the hmmsf program, which is available as part of the HMMER package of search programs, is a family specific default program for MILPAT0063 and a score of 15 is the default threshold score for determining a hit. Alternatively, the threshold score for determining a hit can be lowered (*e.g.*, to 8 bits). A description of the Pfam database can be found in Sonhammer *et al.* (1997) *Proteins* 28(3):405-420 and a detailed description of HMMs can be found, for example, in Gribskov *et al.* (1990) *Meth. Enzymol.* 183:146-159; Gribskov *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:4355-4358; Krogh *et al.* (1994) *J. Mol. Biol.* 235:1501-1531; and Stultz *et al.* (1993) *Protein Sci.* 2:305-314, the contents of which are incorporated herein by reference. A search was performed against the HMM database resulting in the identification of nine ankyrin repeat domains in the amino acid sequence of human CARK (SEQ ID NO:2) at about residues 66-99, 100-132, 133-165, 168-198, 199-233, 234-268, 269-302, 306-338, and 339-371 of SEQ ID NO:2. The results of the search are set forth in Figure 4. Nine ankyrin repeat domains were also identified in the amino acid sequence of rat CARK (SEQ ID NO:8) at about residues 66-99, 100-132, 133-165, 168-198,

199-233, 234-264, 269-302, 306-338, and 339-371 of SEQ ID NO:8. The results of the search are set forth in Figure 6.

In a further preferred embodiment, an ankyrin repeat domain includes at least about 25-40, more preferably about 25-35 amino acid residues, or about 30-35 amino acids and has at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with an ankyrin repeat domain of human or rat CARK (*e.g.*, residues 66-99, 100-132, 133-165, 168-198, 199-233, 234-268, 269-302, 306-338, and 339-371 of SEQ ID NO:2; or residues 66-99, 100-132, 133-165, 168-198, 199-233, 234-264, 269-302, 306-338, and 339-371 of SEQ ID NO:8).

In another embodiment, a CARK of the present invention is identified based on the presence of a "protein kinase domain" in the protein or corresponding nucleic acid molecule. As used herein, the term " protein kinase domain " includes a protein domain having an amino acid sequence of about 200-400 amino acid residues and having a bit score for the alignment of the sequence to the protein kinase domain (HMM) of at least 200. Preferably, a protein kinase domain includes at least about 200-300, and more preferably about 250-300 amino acid residues, and has a bit score for the alignment of the sequence to the protein kinase domain (HMM) of at least 210, 220, 230, 250, 300 or greater. The protein kinase domain (HMM) has been assigned the PFAM Accession PF00069 (<http://genome.wustl.edu/Pfam/.html>).

To identify the presence of a protein kinase domain in a CARK protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of HMMs as described above. A search was performed against the HMM database resulting in the identification of a protein kinase domain in the amino acid sequence of human CARK (SEQ ID NO:2) at about residues 463-716 of SEQ ID NO:2. The results of the search are set forth in Figure 4. A protein kinase domain was also identified in the amino acid sequence of rat CARK (SEQ ID NO:8) at about residues 463-716 of SEQ ID NO:8. The results of the search are set forth in Figure 6.

In another preferred embodiment, a protein kinase domain includes at least about 200-400, more preferably about 200-300 amino acid residues, or about 250-300 amino acids and has at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with a protein kinase domain of human or rat CARK (*e.g.*, residues 463-716 of SEQ ID NO:2 or 8).

Accordingly, CARK proteins having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with an ankyrin repeat domain or a protein kinase domain of human or rat CARK are within the scope of the invention.

In another embodiment, a CARK of the present invention is identified based on the presence of a "LXCXE motif " in the protein or corresponding nucleic acid molecule. The

LXCXE motif is found in mammalian and viral proteins such as cyclin D, E1A, HPV E7, and SV40 large T antigen, and may mediate the interaction with the retinoblastoma protein (pRb) encoded by the retinoblastoma tumor-suppressor gene (RB). The LXCXE motif has the following signature pattern:

5 L-X-C-X-E (SEQ ID NO:10)

A LXCXE motif was identified in the amino acid sequence of human and rat CARK at about residues 411-415 of SEQ ID NO:2 and 8.

Isolated proteins of the present invention, preferably CARK proteins, have an amino acid sequence sufficiently homologous to the amino acid sequence of SEQ ID NO:2 or 8 or are
10 encoded by a nucleotide sequence sufficiently homologous to SEQ ID NO:1, 3, 7, or 9. As used herein, the term "sufficiently homologous" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (*e.g.*, an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or
15 nucleotide sequences share common structural domains or motifs and/or a common functional activity. For example, amino acid or nucleotide sequences which share common structural domains have at least 30%, 40%, or 50% homology, preferably 60% homology, more preferably 70%-80%, and even more preferably 90-95% homology across the amino acid sequences of the domains and contain at least one and preferably two structural domains or
20 motifs, are defined herein as sufficiently homologous. Furthermore, amino acid or nucleotide sequences which share at least 30%, 40%, or 50%, preferably 60%, more preferably 70-80%, or 90-95% homology and share a common functional activity are defined herein as sufficiently homologous.

As used interchangeably herein, a "CARK activity", "biological activity of CARK"
25 or "functional activity of CARK", refers to an activity exerted by a CARK protein, polypeptide or nucleic acid molecule on a CARK responsive cell or on a CARK protein substrate, as determined *in vivo*, or *in vitro*, according to standard techniques. In one embodiment, a CARK activity is a direct activity, such as an association with a CARK-target molecule. As used herein, a "target molecule" or "binding partner" is a molecule with
30 which a CARK protein binds or interacts in nature, such that CARK-mediated function is achieved. A CARK target molecule can be a non-CARK molecule or a CARK protein or polypeptide of the present invention. In an exemplary embodiment, a CARK target molecule is a CARK substrate. Alternatively, a CARK activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the CARK protein with a CARK
35 ligand or substrate. The biological activity of CARK is described herein.

In a preferred embodiment, a CARK activity is at least one of the following activities: (i) interaction with a CARK substrate or target molecule (*e.g.*, a CARK polypeptide or a non-CARK molecule (*e.g.*, pRb)); (ii) transfer of a phosphate group to a

CARK substrate or target molecule; (iii) modulation of intra- or inter-cellular signaling and/or gene transcription; and (iv) modulation of cell proliferation, growth, differentiation, survival and/or migration. In another preferred embodiment, a CARK activity is a dual kinase activity, *e.g.*, the transfer of a phosphate group to serine, threonine and tyrosine residues.

Accordingly, another embodiment of the invention features isolated CARK proteins and polypeptides having a CARK activity. Preferred proteins are CARK proteins having at least one ankyrin repeat domain and preferably two, three, four, five, six, seven, eight, or, most preferably, nine or more ankyrin repeat domains, and, preferably, a CARK activity.

Other preferred proteins are CARK proteins having a protein kinase domain and, preferably, a CARK activity. Yet other preferred proteins have at least one ankyrin repeat domain and preferably two, three, four, five, six, seven, eight, or, most preferably, nine or more ankyrin repeat domains, a protein kinase domain, and, preferably, a CARK activity. Additional preferred proteins have at least one ankyrin repeat domain and preferably two, three, four, five, six, seven, eight, or, most preferably, nine or more ankyrin repeat domains, a protein kinase domain, and are, preferably, encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 7, or 9.

The nucleotide sequence of the isolated human CARK cDNA and the predicted amino acid sequence of the human CARK polypeptide are shown in Figure 1 and in SEQ ID NOs:1 and 2, respectively. A plasmid containing the nucleotide sequence encoding human CARK was deposited with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on _____ and assigned Accession Number _____. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

The human CARK gene, which is approximately 3025 nucleotides in length, encodes a protein having a molecular weight of approximately 96 kD and which is approximately 835 amino acid residues in length.

The nucleotide sequence of the isolated rat CARK cDNA and the predicted amino acid sequence of the rat CARK polypeptide are shown in Figure 5 and in SEQ ID NOs:7 and 8, respectively. A plasmid containing the nucleotide sequence encoding rat CARK was deposited with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on _____ and assigned Accession Number _____. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This

deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

The rat CARK gene, which is approximately 3026 nucleotides in length, encodes a protein having a molecular weight of approximately 96 kD and which is approximately 835 amino acid residues in length.

Various aspects of the invention are described in further detail in the following subsections:

10 I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode CARK proteins or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes to identify CARK-encoding nucleic acid molecules (*e.g.*, CARK mRNA) and fragments for use as PCR primers for the amplification or mutation of CARK nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

The term "isolated nucleic acid molecule" includes nucleic acid molecules which are separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example, with regards to genomic DNA, the term "isolated" includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated CARK nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or portion of the nucleic acid sequence of

SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, as a hybridization probe, CARK nucleic acid molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, 5 Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ can be isolated by the polymerase chain reaction (PCR) using 10 synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers 15 according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to CARK nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In a preferred embodiment, an isolated nucleic acid molecule of the invention 20 comprises the nucleotide sequence shown in SEQ ID NO:1. The sequence of SEQ ID NO:1 corresponds to the human CARK cDNA. This cDNA comprises sequences encoding the human CARK protein (*i.e.*, "the coding region", from nucleotides 48-2552), as well as 5' untranslated sequences (nucleotides 1-47) and 3' untranslated sequences (nucleotides 2553-3025). Alternatively, the nucleic acid molecule can comprise only the coding region of SEQ 25 ID NO:1 (*e.g.*, nucleotides 48-2552, corresponding to SEQ ID NO:3).

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO:7. The sequence of SEQ ID NO:7 corresponds to the rat CARK cDNA. This cDNA comprises sequences encoding the rat CARK protein (*i.e.*, "the coding region", from nucleotides 61-2565), as well as 5' 30 untranslated sequences (nucleotides 1-60) and 3' untranslated sequences (nucleotides 2566-3026). Alternatively, the nucleic acid molecule can comprise only the coding region of SEQ ID NO:7 (*e.g.*, nucleotides 61-2565, corresponding to SEQ ID NO:9).

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence shown 35 in SEQ ID NO:1, 3, 7, or 9; or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1, 3, 7, or 9 or the nucleotide sequence of the DNA insert of the

plasmid deposited with ATCC as Accession Number _____, is one which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that it can hybridize to the nucleotide sequence shown in SEQ ID
5 NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 82%, 85%, 90%, 95%, 98% or more homologous to the entire length
10 of the nucleotide sequence shown in SEQ ID NO:1, 3, 7, or 9, or the entire length of the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion of any of these nucleotide sequences.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the nucleic acid sequence of SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the
15 DNA insert of the plasmid deposited with ATCC as Accession Number _____, for example, a fragment which can be used as a probe or primer or a fragment encoding a portion of a CARK protein, *e.g.*, a biologically active portion of a CARK protein. The nucleotide sequence determined from the cloning of the CARK gene allows for the generation of probes and primers designed for use in identifying and/or cloning other CARK family
20 members, as well as CARK homologues from other species. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 consecutive nucleotides of a sense sequence of SEQ ID NO:1, 3, 7, or 9, or the nucleotide
25 sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, of an anti-sense sequence of SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number
30 _____. In an exemplary embodiment, a nucleic acid molecule of the present invention comprises a nucleotide sequence which is greater than 250-300, 300-350, 350-400, 400-450, 467, 467-500, 500-550, or 550-600, 600-800, 800-1000, 1000-1200, 1200-1400, 1400-1600, 1600-1800, 1800-2000, 2000-2400, 2400-2800, 2800-2900, 2962 or more nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of
35 SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

Probes based on the CARK nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments,

the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a CARK protein, such as by measuring a level of a CARK-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting CARK mRNA levels or determining whether a genomic CARK gene has been mutated or deleted.

A nucleic acid fragment encoding a "biologically active portion of a CARK protein" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, which encodes a polypeptide having a CARK biological activity (the biological activities of the CARK proteins are described herein), expressing the encoded portion of the CARK protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the CARK protein.

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, due to degeneracy of the genetic code and thus encode the same CARK proteins as those encoded by the nucleotide sequence shown in SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO:2 or 8.

In addition to the CARK nucleotide sequences shown in SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the CARK proteins may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the CARK genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which include an open reading frame encoding a CARK protein, preferably a mammalian CARK protein, and can further include non-coding regulatory sequences, and introns.

Allelic variants of human and rat CARK include both functional and non-functional CARK proteins. Functional allelic variants are naturally occurring amino acid sequence variants of the human and rat CARK proteins that maintain the ability to bind a CARK ligand and/or modulate cellular mechanisms associated with cell growth or differentiation. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2 or 8 or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein.

Non-functional allelic variants are naturally occurring amino acid sequence variants of the human and rat CARK proteins that do not have the ability to either bind a CARK ligand and/or modulate cellular mechanisms associated with cell growth or differentiation. Non-functional allelic variants will typically contain a non-conservative substitution, a
5 deletion, or insertion or premature truncation of the amino acid sequence of SEQ ID NO:2 or 8, or a substitution, insertion or deletion in critical residues or critical regions.

The present invention further provides non-human orthologues of the human CARK protein. Orthologues of the human CARK protein are proteins that are isolated from non-human organisms and possess the same CARK ligand and/or substrate binding, and/or
10 modulation of cellular mechanisms associated with cell growth or differentiation of the human CARK protein, *e.g.* the rat CARK protein (SEQ ID NO:8). Orthologues of the human CARK protein can readily be identified as comprising an amino acid sequence that is substantially homologous to SEQ ID NO:2 or 8.

Moreover, nucleic acid molecules encoding other CARK family members and, thus,
15 which have a nucleotide sequence which differs from the CARK sequences of SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention. For example, another CARK cDNA can be identified based on the nucleotide sequence of human or rat CARK. Moreover, nucleic acid molecules encoding CARK
20 proteins from different species, and which, thus, have a nucleotide sequence which differs from the CARK sequences of SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention. For example, a mouse CARK cDNA can be identified based on the nucleotide sequence of a human or rat CARK.

25 Nucleic acid molecules corresponding to natural allelic variants and homologues of the CARK cDNAs of the invention can be isolated based on their homology to the CARK nucleic acids disclosed herein using the cDNAs disclosed herein, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Nucleic acid molecules corresponding to natural allelic variants
30 and homologues of the CARK cDNAs of the invention can further be isolated by mapping to the same chromosome or locus as the CARK gene.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 15, 20, 25, 30 or more nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ
35 ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In other embodiment, the nucleic acid is at least 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 467, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1500, 2000, 2500, or 2962 or more nucleotides in length.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% identical to each other remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4, and 6. Additional stringent conditions can be found in *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9, and 11. A preferred, non-limiting example of stringent hybridization conditions includes hybridization in 4X sodium chloride/sodium citrate (SSC), at about 65-70°C (or alternatively hybridization in 4X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70°C. A preferred, non-limiting example of highly stringent hybridization conditions includes hybridization in 1X SSC, at about 65-70°C (or alternatively hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions includes hybridization in 4X SSC, at about 50-60°C (or alternatively hybridization in 6X SSC plus 50% formamide at about 40-45°C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, *e.g.*, at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1xSSPE is 0.15M NaCl, 10mM NaH₂PO₄, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1X SSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, $T_m(^{\circ}\text{C}) = 2(\# \text{ of A} + \text{T bases}) + 4(\# \text{ of G} + \text{C bases})$. For hybrids between 18 and 49 base pairs in length, $T_m(^{\circ}\text{C}) = 81.5 + 16.6(\log_{10}[\text{Na}^+]) + 0.41(\% \text{G} + \text{C}) - (600/\text{N})$, where N is the number of bases in the hybrid, and $[\text{Na}^+]$ is the concentration of sodium ions in the hybridization buffer ($[\text{Na}^+]$ for 1X SSC = 0.165 M). It will also be recognized by the skilled practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (*e.g.*, BSA or salmon or herring sperm carrier DNA), detergents (*e.g.*, SDS), chelating agents (*e.g.*, EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH₂PO₄, 7% SDS at about 65°C, followed by one

or more washes at 0.02M NaH₂PO₄, 1% SDS at 65°C (see *e.g.*, Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995), or alternatively 0.2X SSC, 1% SDS.

Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1, 3, 7, or 9, corresponds to a naturally-
 5 occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In addition to naturally-occurring allelic variants of the CARK sequences that may exist in the population, the skilled artisan will further appreciate that changes can be
 10 introduced by mutation into the nucleotide sequences of SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby leading to changes in the amino acid sequence of the encoded CARK proteins, without altering the functional ability of the CARK proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid
 15 residues can be made in the sequence of SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of CARK (*e.g.*, the sequence of SEQ ID NO:2 or 8) without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For
 20 example, amino acid residues that are conserved among the CARK proteins of the present invention, *e.g.*, those present in the active site of the protein kinase domain, are predicted to be particularly unamenable to alteration. Furthermore, additional amino acid residues that are conserved between the CARK proteins of the present invention and other ankyrin repeat containing kinases are not likely to be amenable to alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules
 25 encoding CARK proteins that contain changes in amino acid residues that are not essential for activity. Such CARK proteins differ in amino acid sequence from SEQ ID NO:2 or 8, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid
 30 sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 92%, 95%, 98% or more homologous to SEQ ID NO:2 or 8.

An isolated nucleic acid molecule encoding a CARK protein homologous to the protein of SEQ ID NO:2 or 8 can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1, 3, 7, or 9,
 35 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with

ATCC as Accession Number _____ by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a CARK protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a CARK coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for CARK biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

In one embodiment, a mutant CARK polypeptide comprises an amino acid substitution at an amino acid residue in the protein kinase domain, *e.g.*, lysine residue 490 of SEQ ID NO:2 or 8. In another embodiment, a mutant CARK polypeptide comprises an amino acid substitution and/or deletion in a N-myristoylation site, *e.g.*, residues 2-7 of SEQ ID NO:2. In yet another embodiment, a mutant CARK polypeptide comprises an amino acid substitution and/or deletion mutation in a LXCXE motif, *e.g.*, residues 411-415 of SEQ ID NO:2.

In a preferred embodiment, a mutant CARK protein can be assayed for the ability to 1) regulate transmission of signals from cellular receptors, *e.g.*, cardiac cell growth factor receptors; 2) modulate the entry of cells, *e.g.*, cardiac precursor cells, into mitosis; 3) modulate cellular differentiation; 4) modulate cell death; and 5) regulate cytoskeleton function, *e.g.*, actin bundling.

In addition to the nucleic acid molecules encoding CARK proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire CARK coding strand, or to

only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding CARK. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the coding region of human CARK corresponds to SEQ ID NO:3, and the coding region of rat CARK corresponds to SEQ ID NO:9). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding CARK. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding CARK disclosed herein (*e.g.*, SEQ ID NO:3 or 9), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of CARK mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of CARK mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of CARK mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted

nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave CARK mRNA transcripts to thereby inhibit translation of CARK mRNA. A ribozyme having specificity for a CARK-encoding nucleic acid can be designed based upon the nucleotide sequence of a CARK cDNA disclosed herein (*i.e.*, SEQ ID NO:1, 3, 7, or 9, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a CARK-encoding mRNA. See, *e.g.*, Cech *et al.* U.S.

Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, CARK mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

5 Alternatively, CARK gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the CARK gene (*e.g.*, the CARK promoter and/or enhancers) to form triple helical structures that prevent transcription of the CARK gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. 10 (1992) *Bioassays* 14(12):807-15.

In yet another embodiment, the CARK nucleic acid molecules of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate 15 peptide nucleic acids (see Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA 20 under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *et al.* *Proc. Natl. Acad. Sci.* 93: 14670-675.

PNAs of CARK nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation 25 arrest or inhibiting replication. PNAs of CARK nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (*e.g.*, by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (*e.g.*, S1 nucleases (Hyrup B. (1996) *supra*)); or as probes or primers for DNA sequencing or 30 hybridization (Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *supra*).

In another embodiment, PNAs of CARK can be modified, (*e.g.*, to enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of CARK nucleic acid 35 molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (*e.g.*, RNase H and DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of

appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup B. (1996) *supra* and Finn P.J. *et al.* (1996) *Nucleic Acids Res.* 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag, M. *et al.* (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn P.J. *et al.* (1996) *supra*).
Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, *e.g.*, Krol *et al.* (1988) *Bio-Techniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (*e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

II. Isolated CARK Proteins and Anti-CARK Antibodies

One aspect of the invention pertains to isolated CARK proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-CARK antibodies. In one embodiment, native CARK proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, CARK proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a CARK protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the CARK protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of CARK protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of CARK protein having less than about 30% (by dry weight) of non-CARK

protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-CARK protein, still more preferably less than about 10% of non-CARK protein, and most preferably less than about 5% non-CARK protein. When the CARK protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of CARK protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of CARK protein having less than about 30% (by dry weight) of chemical precursors or non-CARK chemicals, more preferably less than about 20% chemical precursors or non-CARK chemicals, still more preferably less than about 10% chemical precursors or non-CARK chemicals, and most preferably less than about 5% chemical precursors or non-CARK chemicals.

As used herein, a "biologically active portion" of a CARK protein includes a fragment of a CARK protein which participates in an interaction between a CARK molecule and a non-CARK molecule. Biologically active portions of a CARK protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the CARK protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2 or 8, which include less amino acids than the full length CARK proteins, and exhibit at least one activity of a CARK protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the CARK protein, *e.g.*, modulating signaling pathways associated with cellular growth and differentiation. A biologically active portion of a CARK protein can be a polypeptide which is, for example, 10, 25, 50, 100, 200, 250, 254, 300, 306, 350, 400, 412, 428, 450 or more amino acids in length. Biologically active portions of a CARK protein can be used as targets for developing agents which modulate a CARK mediated activity, *e.g.*, the modulation of signaling pathways associated with cellular growth and differentiation.

In one embodiment, a biologically active portion of a CARK protein comprises at least one ankyrin repeat domain, and/or at least one protein kinase domain. In another embodiment, a biologically active portion of a CARK protein comprises at least one of the following domains and/or motifs: an ankyrin repeat domain, a kinase domain or a LXCXE motif. It is to be understood that a preferred biologically active portion of a CARK protein of the present invention may contain at least one ankyrin repeat domain. Another preferred biologically active portion of a CARK protein may contain at least two, three, four, five, six, seven, eight, or nine ankyrin repeat domains. Yet another preferred biologically active

portion of a CARK protein may contain at least one protein kinase domain. Another preferred biologically active portion of a CARK protein may contain at least one protein kinase domain and at least one LXCXE motif. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native CARK protein.

In a preferred embodiment, the CARK protein has an amino acid sequence shown in SEQ ID NO:2 or 8. In other embodiments, the CARK protein is substantially homologous to SEQ ID NO:2 or 8, and retains the functional activity of the protein of SEQ ID NO:2 or 8, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the CARK protein is a protein which comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to SEQ ID NO:2 or 8.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (*e.g.*, when aligning a second sequence to the CARK amino acid sequence of SEQ ID NO:2 having 835 amino acid residues, at least 251, preferably at least 334, more preferably at least 418, even more preferably at least 501, and even more preferably at least 585, 668 or 752 amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent

identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (Comput Appl Biosci, 4:11-17 (1988)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to CARK nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to CARK protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.

The invention also provides CARK chimeric or fusion proteins. As used herein, a CARK "chimeric protein" or "fusion protein" comprises a CARK polypeptide operatively linked to a non-CARK polypeptide. An "CARK polypeptide" refers to a polypeptide having an amino acid sequence corresponding to CARK, whereas a "non-CARK polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the CARK protein, *e.g.*, a protein which is different from the CARK protein and which is derived from the same or a different organism. Within a CARK fusion protein the CARK polypeptide can correspond to all or a portion of a CARK protein. In a preferred embodiment, a CARK fusion protein comprises at least one biologically active portion of a CARK protein. In another preferred embodiment, a CARK fusion protein comprises at least two biologically active portions of a CARK protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the CARK polypeptide and the non-CARK polypeptide are fused in-frame to each other. The non-CARK polypeptide can be fused to the N-terminus or C-terminus of the CARK polypeptide.

For example, in one embodiment, the fusion protein is a GST-CARK fusion protein in which the CARK sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant CARK. In another embodiment, the fusion protein is a GFP-CARK fusion protein. Such fusion proteins can

facilitate the visualization of CARK expression (*e.g.*, visualization of subcellular localization). In further embodiments, CARK sequences can be fused to poly-histidine residues, or an amino acid sequence epitope derived from a heterologous protein, *e.g.*, myc.

In another embodiment, the fusion protein is a CARK protein containing a
5 heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of CARK can be increased through use of a heterologous signal sequence.

The CARK fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The CARK fusion proteins can be used
10 to affect the bioavailability of a CARK substrate. Use of CARK fusion proteins may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a CARK protein; (ii) mis-regulation of the CARK gene; and (iii) aberrant post-translational modification of a CARK protein.

Moreover, the CARK-fusion proteins of the invention can be used as immunogens to
15 produce anti-CARK antibodies in a subject, to purify CARK ligands and in screening assays to identify molecules which inhibit the interaction of CARK with a CARK substrate.

Preferably, a CARK chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with
20 conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR
25 amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion
30 moiety (*e.g.*, a GST polypeptide). A CARK-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the CARK protein.

The present invention also pertains to variants of the CARK proteins which function as either CARK agonists (mimetics) or as CARK antagonists. Variants of the CARK proteins can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of a
35 CARK protein. An agonist of the CARK proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a CARK protein. An antagonist of a CARK protein can inhibit one or more of the activities of the naturally occurring form of the CARK protein by, for example, competitively modulating a CARK-

mediated activity of a CARK protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the CARK protein.

In one embodiment, a variant of a CARK protein functions as a dominant negative CARK molecule which antagonizes a CARK activity. In a preferred embodiment, a dominant negative CARK molecule contains an amino acid substitution, insertion and/or deletion in the protein kinase domain, *e.g.*, a lysine to arginine substitution at amino acid residue 490 of SEQ ID NO:2 or 8. In another embodiment, a variant of a CARK protein functions as a hyperactive CARK molecule which acts as a CARK agonist. In a preferred embodiment, an agonist of a CARK molecule contains an amino acid substitution, insertion and/or deletion in the protein kinase domain.

In one embodiment, variants of a CARK protein which function as either CARK agonists (mimetics) or as CARK antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of a CARK protein for CARK protein agonist or antagonist activity. In one embodiment, a variegated library of CARK variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of CARK variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential CARK sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of CARK sequences therein. There are a variety of methods which can be used to produce libraries of potential CARK variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential CARK sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477).

In addition, libraries of fragments of a CARK protein coding sequence can be used to generate a variegated population of CARK fragments for screening and subsequent selection of variants of a CARK protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a CARK coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single

stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the CARK protein.

5 Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of CARK proteins. The most widely used techniques, which are amenable to high through-put analysis, for screening
10 large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in
15 the libraries, can be used in combination with the screening assays to identify CARK variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

 In one embodiment, cell based assays can be exploited to analyze a variegated CARK library. For example, a library of expression vectors can be transfected into a cell
20 line, *e.g.*, a cardiac cell line, which ordinarily responds to a particular ligand in a CARK-dependent manner. The transfected cells are then contacted with the ligand and the effect of expression of the mutant on signaling by the ligand can be detected, *e.g.*, by monitoring intracellular calcium, IP3, or diacylglycerol concentration, phosphorylation profile of intracellular proteins, cell proliferation and/or migration, or the activity of a CARK-
25 regulated transcription factor. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of signaling by the ligand, and the individual clones further characterized.

 An isolated CARK protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind CARK using standard techniques for polyclonal
30 and monoclonal antibody preparation. A full-length CARK protein can be used or, alternatively, the invention provides antigenic peptide fragments of CARK for use as immunogens. The antigenic peptide of CARK comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2 or 8 and encompasses an epitope of CARK such that an antibody raised against the peptide forms a specific immune complex with
35 CARK. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues. In one embodiment, the

antigenic peptide comprises the C-terminal portion of a CARK polypeptide, *e.g.*, amino acid residues 796-835 of SEQ ID NO:2 or 8.

Preferred epitopes encompassed by the antigenic peptide are regions of CARK that are located on the surface of the protein, *e.g.*, hydrophilic regions, as well as regions with high antigenicity (see, for example, Figure 2).

A CARK immunogen typically is used to prepare antibodies by immunizing a suitable subject, (*e.g.*, rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed CARK protein or a chemically synthesized CARK polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic CARK preparation induces a polyclonal anti-CARK antibody response.

Accordingly, another aspect of the invention pertains to anti-CARK antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as CARK. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')₂ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind CARK. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of CARK. A monoclonal antibody composition thus typically displays a single binding affinity for a particular CARK protein with which it immunoreacts. In one embodiment, a monoclonal antibody of the invention binds specifically to an epitope comprised in the C-terminus of a CARK polypeptide, *e.g.*, amino acid residues 796-835 of SEQ ID NO:2 or 8

Polyclonal anti-CARK antibodies can be prepared as described above by immunizing a suitable subject with a CARK immunogen. The anti-CARK antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized CARK. If desired, the antibody molecules directed against CARK can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, *e.g.*, when the anti-CARK antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497) (see also, Brown *et al.* (1981) *J. Immunol.* 127:539-46; Brown *et al.* (1980) *J. Biol. Chem.* 255:4980-83; Yeh *et al.* (1976) *Proc. Natl. Acad. Sci. USA* 76:2927-31; and Yeh *et al.*

(1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see generally R. H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); E. A. Lerner (1981) *Yale J. Biol. Med.*, 54:387-402; M. L. Gefter *et al.* (1977) *Somatic Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a CARK immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds CARK.

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-CARK monoclonal antibody (see, *e.g.*, G. Galfre *et al.* (1977) *Nature* 266:55052; Gefter *et al.* *Somatic Cell Genet.*, cited *supra*; Lerner, *Yale J. Biol. Med.*, cited *supra*; Kenneth, *Monoclonal Antibodies*, cited *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (*e.g.*, a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, *e.g.*, the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind CARK, *e.g.*, using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-CARK antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with CARK to thereby isolate immunoglobulin library members that bind CARK. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene

SurfZAP™ Phage Display Kit, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* PCT International Publication No. WO 92/18619; Dower *et al.* PCT International Publication No. WO 91/17271; Winter *et al.* PCT International Publication WO 92/20791; Markland *et al.* PCT International Publication No. WO 92/15679; Breitling *et al.* PCT International Publication WO 93/01288; McCafferty *et al.* PCT International Publication No. WO 92/01047; Garrard *et al.* PCT International Publication No. WO 92/09690; Ladner *et al.* PCT International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrad *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc. Acid Res.* 19:4133-4137; Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty *et al.* *Nature* (1990) 348:552-554.

Additionally, recombinant anti-CARK antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson *et al.* International Application No. PCT/US86/02269; Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT International Publication No. WO 86/01533; Cabilly *et al.* U.S. Patent No. 4,816,567; Cabilly *et al.* European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559; Morrison, S. L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

An anti-CARK antibody (*e.g.*, monoclonal antibody) can be used to isolate CARK by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-CARK antibody can facilitate the purification of natural CARK from cells and of recombinantly produced CARK expressed in host cells. Moreover, an anti-CARK antibody can be used to detect CARK protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the CARK protein. Anti-CARK

antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a CARK protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions. In one embodiment, a CARK polypeptide is expressed in a host cell using a Herpes Simplex Virus (HSV) vector.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected

on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990).

Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., CARK proteins, mutant forms of CARK proteins, fusion proteins, and the like).

The recombinant expression vectors of the invention can be designed for expression of CARK proteins in prokaryotic or eukaryotic cells. For example, CARK proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5

(Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Purified fusion proteins can be utilized in CARK activity assays, (e.g., direct assays or competitive assays described in detail below), or to generate antibodies specific for CARK proteins, for example. In a preferred embodiment, a CARK fusion protein expressed in a retroviral expression vector of the present invention can be utilized to infect bone marrow cells which are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (e.g., six (6) weeks).

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the CARK expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and picZ (Invitrogen Corp, San Diego, CA).

Alternatively, CARK proteins can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC

(Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic
5 cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.*

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory
10 elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the alpha-myosin heavy chain promoter (cardiac specific); albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors
15 (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S.
20 Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α -fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The expression characteristics of an endogenous CARK gene within a cell line or
25 microorganism may be modified by inserting a heterologous DNA regulatory element into the genome of a stable cell line or cloned microorganism such that the inserted regulatory element is operatively linked with the endogenous CARK gene. For example, an endogenous CARK gene which is normally "transcriptionally silent", *i.e.*, a CARK gene which is normally not expressed, or is expressed only at very low levels in a cell line or
30 microorganism, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell line or microorganism. Alternatively, a transcriptionally silent, endogenous CARK gene may be activated by insertion of a promiscuous regulatory element that works across cell types.

A heterologous regulatory element may be inserted into a stable cell line or cloned
35 microorganism, such that it is operatively linked with an endogenous CARK gene, using techniques, such as targeted homologous recombination, which are well known to those of skill in the art, and described, *e.g.*, in Chappel, U.S. Patent No. 5,272,071; PCT publication No. WO 91/06667, published May 16, 1991.

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to CARK mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a CARK nucleic acid molecule of the invention is introduced, *e.g.*, a CARK nucleic acid molecule within a recombinant expression vector or a CARK nucleic acid molecule containing sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, a CARK protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Vero cells, cardiac myocytes, Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual*, 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding a CARK protein or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) a CARK protein. Accordingly, the invention further provides methods for producing a CARK protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector encoding a CARK protein has been introduced) in a suitable medium such that a CARK protein is produced. In another embodiment, the method further comprises isolating a CARK protein from the medium or the host cell.

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which CARK-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous CARK sequences have been introduced into their genome or homologous recombinant animals in which endogenous CARK sequences have been altered. Such animals are useful for studying the function and/or activity of a CARK and for identifying and/or evaluating modulators of CARK activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous CARK gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing a CARK-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The CARK cDNA sequence of SEQ ID NO:1 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human CARK gene, such as a mouse or rat CARK gene (*e.g.*, SEQ ID NO:7), can be used as a transgene. Alternatively, a CARK gene homologue, such as another CARK family member, can be isolated based on hybridization to the CARK cDNA sequences of SEQ ID NO:1, 3, 7, or 9, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____ (described further in subsection I above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a CARK transgene to direct expression of a CARK protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of a CARK transgene in its genome and/or expression of CARK mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a CARK protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a CARK gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the CARK gene. The CARK gene can be a human gene (*e.g.*, the cDNA of SEQ ID NO:3), but more preferably, is a non-human homologue of a human CARK gene (*e.g.*, SEQ ID NO:9, or a cDNA isolated by stringent hybridization with the nucleotide sequence of SEQ ID NO:1 or 7). For example, a mouse CARK gene can be used to construct a homologous recombination nucleic acid molecule, *e.g.*, a vector, suitable for altering an endogenous CARK gene in the mouse genome. In a preferred embodiment, the homologous recombination nucleic acid molecule is designed such that, upon homologous recombination, the endogenous CARK gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the homologous recombination nucleic acid molecule can be designed such that, upon homologous recombination, the endogenous CARK gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream

regulatory region can be altered to thereby alter the expression of the endogenous CARK protein). In the homologous recombination nucleic acid molecule, the altered portion of the CARK gene is flanked at its 5' and 3' ends by additional nucleic acid sequence of the CARK gene to allow for homologous recombination to occur between the exogenous CARK gene carried by the homologous recombination nucleic acid molecule and an endogenous CARK gene in a cell, *e.g.*, an embryonic stem cell. The additional flanking CARK nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the homologous recombination nucleic acid molecule (see, *e.g.*, Thomas, K.R. and Capecchi, M. R. (1987) *Cell* 51:503 for a description of homologous recombination vectors). The homologous recombination nucleic acid molecule is introduced into a cell, *e.g.*, an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced CARK gene has homologously recombined with the endogenous CARK gene are selected (see *e.g.*, Li, E. *et al.* (1992) *Cell* 69:915). The selected cells can then be injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras (see *e.g.*, Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination nucleic acid molecules, *e.g.*, vectors, or homologous recombinant animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec *et al.*; WO 91/01140 by Smithies *et al.*; WO 92/0968 by Zijlstra *et al.*; and WO 93/04169 by Berns *et al.*

In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al.* (1997) *Nature* 385:810-813 and PCT

International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

10 IV. Pharmaceutical Compositions

The CARK nucleic acid molecules, fragments of CARK proteins, and anti-CARK antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration,

suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a fragment of a CARK protein or an anti-CARK antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such

as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically

effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

As defined herein, a therapeutically effective amount of protein or polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The skilled artisan will appreciate that certain factors may influence the dosage required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

In a preferred example, a subject is treated with antibody, protein, or polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody, protein, or polypeptide used for treatment may increase or decrease over the course of a particular treatment. Changes in dosage may result and become apparent from the results of diagnostic assays as described herein.

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or

inorganic compounds (*i.e.*, including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds. It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention.

Exemplary doses include milligram or microgram amounts of the small molecule per kilogram of subject or sample weight (*e.g.*, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram. It is furthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

Further, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, teniposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine

(CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, α -interferon, β -interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon *et al.*, "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in *Monoclonal Antibodies And Cancer Therapy*, Reisfeld *et al.* (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom *et al.*, "Antibodies For Drug Delivery", in *Controlled Drug Delivery* (2nd Ed.), Robinson *et al.* (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in *Monoclonal Antibodies '84: Biological And Clinical Applications*, Pinchera *et al.* (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in *Monoclonal Antibodies For Cancer Detection And Therapy*, Baldwin *et al.* (eds.), pp. 303-16 (Academic Press 1985), and Thorpe *et al.*, "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", *Immunol. Rev.*, 62:119-58 (1982). Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by stereotactic injection (see *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

V. Uses and Methods of the Invention

5 The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) predictive medicine (*e.g.*, diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and c) methods of treatment (*e.g.*, therapeutic and prophylactic). As described herein, a CARK protein of the invention has one or more of the following
10 activities: 1) it regulates transmission of signals from cellular receptors, *e.g.*, cardiac cell growth factor receptors; 2) it modulates the entry of cells, *e.g.*, cardiac precursor cells, into mitosis; 3) it modulates cellular differentiation; 4) it modulates cell death; and 5) it regulates cytoskeleton function, *e.g.*, actin bundling, and, thus, can be used to, for example, 1) regulate transmission of signals from cellular receptors, *e.g.*, cardiac cell growth factor
15 receptors; 2) modulate the entry of cells, *e.g.*, cardiac precursor cells, into mitosis; 3) modulate cellular differentiation; 4) modulate cell death; and 5) regulate cytoskeleton function, *e.g.*, actin bundling.

The isolated nucleic acid molecules of the invention can be used, for example, to express CARK protein (*e.g.*, via a recombinant expression vector in a host cell in gene
20 therapy applications), to detect CARK mRNA (*e.g.*, in a biological sample) or a genetic alteration in a CARK gene, and to modulate CARK activity, as described further below. The CARK proteins can be used to treat disorders characterized by insufficient or excessive production of a CARK substrate or production of CARK inhibitors. In addition, the CARK proteins can be used to screen for naturally occurring CARK substrates, to screen for drugs
25 or compounds which modulate CARK activity, as well as to treat disorders characterized by insufficient or excessive production of CARK protein or production of CARK protein forms which have decreased, aberrant or unwanted activity compared to CARK wild type protein (*e.g.*, cellular proliferation or differentiation associated disorders). Moreover, the anti-CARK antibodies of the invention can be used to detect and isolate CARK proteins, regulate
30 the bioavailability of CARK proteins, and modulate CARK activity.

A. Screening Assays:

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides,
35 peptidomimetics, small molecules or other drugs) which bind to CARK proteins, have a stimulatory or inhibitory effect on, for example, CARK expression or CARK activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of CARK substrate.

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a CARK protein or polypeptide or biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a CARK protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994) *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner U.S. Patent 5,223,409), spores (Ladner U.S. Patent '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a CARK protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate CARK activity is determined. Determining the ability of the test compound to modulate CARK activity can be accomplished by monitoring, for example, intracellular calcium, IP3, or diacylglycerol concentration, phosphorylation profile of intracellular proteins, cell proliferation and/or migration, *de novo* protein synthesis, binding of a CARK target molecule, or the activity of a CARK-regulated transcription factor. The cell, for example, can be of mammalian origin, *e.g.*, a cardiac cell.

The ability of the test compound to modulate CARK binding to a substrate or to bind to CARK can also be determined. Determining the ability of the test compound to modulate CARK binding to a substrate can be accomplished, for example, by coupling the CARK substrate with a radioisotope or enzymatic label such that binding of the CARK substrate to CARK can be determined by detecting the labeled CARK substrate in a complex.

Determining the ability of the test compound to bind CARK can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to CARK can be determined by detecting the labeled CARK compound in a complex. For example, compounds (*e.g.*, CARK substrates) can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

It is also within the scope of this invention to determine the ability of a compound (*e.g.*, a CARK substrate) to interact with CARK without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with CARK without the labeling of either the compound or the CARK.

McConnell, H. M. *et al.* (1992) *Science* 257:1906-1912. As used herein, a

"microphysiometer" (*e.g.*, Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a compound and CARK.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a CARK target molecule (*e.g.*, a CARK substrate) with a test compound and determining the ability of the test compound to modulate (*e.g.* stimulate or inhibit) the activity of the CARK target molecule. Determining the ability of the test compound to modulate the activity of a CARK target molecule can be accomplished, for example, by determining the ability of the CARK protein to bind to or interact with the CARK target molecule.

Determining the ability of the CARK protein or a biologically active fragment thereof, to bind to or interact with a CARK target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of the CARK protein to bind to or interact with a CARK target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.*, intracellular Ca^{2+} , diacylglycerol, IP_3 , and the like), detecting catalytic/enzymatic activity of the target on an appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a target-regulated cellular response.

In a preferred embodiment, determining the ability of the test compound to modulate the activity of a CARK target molecule can be accomplished, for example, by determining

the ability of the CARK protein or polypeptide to phosphorylate the target molecule. The ability of the CARK protein or polypeptide to phosphorylate a target molecule can be determined by, for example, a kinase assay, as exemplified herein. In one embodiment, the kinase activity of a CARK protein or polypeptide (*e.g.*, myc-CARK) can be assessed by monitoring autophosphorylation of the CARK protein or polypeptide.

In yet another embodiment, an assay of the present invention is a cell-free assay in which a CARK protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the CARK protein or biologically active portion thereof is determined. Preferred biologically active portions of the CARK proteins to be used in assays of the present invention include fragments which participate in interactions with non-CARK molecules, *e.g.*, fragments with high surface probability scores (see, for example, Figure 2). Binding of the test compound to the CARK protein can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the CARK protein or biologically active portion thereof with a known compound which binds CARK to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a CARK protein, wherein determining the ability of the test compound to interact with a CARK protein comprises determining the ability of the test compound to preferentially bind to CARK or biologically active portion thereof as compared to the known compound.

In another embodiment, the assay is a cell-free assay in which a CARK protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the CARK protein or biologically active portion thereof is determined. In a preferred embodiment, determining the ability of the test compound to modulate the activity of a CARK protein can be accomplished, for example, by determining the ability of the CARK protein or polypeptide to phosphorylate a target molecule (*e.g.*, a CARK substrate). The ability of the CARK protein or polypeptide to phosphorylate a target molecule can be determined by, for example, an *in vitro* kinase assay, as exemplified herein. In one embodiment, the kinase activity of a CARK protein or polypeptide (*e.g.*, GST-CARK or GST-CARK-KD) can be assessed by monitoring autophosphorylation of the CARK protein or polypeptide. In another embodiment, the kinase activity of a CARK protein or polypeptide can be assessed by monitoring phosphorylation of a heterologous substrate, *e.g.*, H1 histone, myelin basic protein, ATF-2 and Phas-1.

Determining the ability of the test compound to modulate the activity of a CARK protein can also be accomplished, for example, by determining the ability of the CARK protein to bind to a CARK target molecule by one of the methods described above for determining direct binding. Determining the ability of the CARK protein to bind to a CARK target molecule can also be accomplished using a technology such as real-time

Biomolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In an alternative embodiment, determining the ability of the test compound to modulate the activity of a CARK protein can be accomplished by determining the ability of the CARK protein to further modulate the activity of a downstream effector of a CARK target molecule. For example, the activity of the effector molecule on an appropriate target can be determined or the binding of the effector to an appropriate target can be determined, as previously described.

In yet another embodiment, the cell-free assay involves contacting a CARK protein or biologically active portion thereof with a known compound which binds the CARK protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the CARK protein, wherein determining the ability of the test compound to interact with the CARK protein comprises determining the ability of the CARK protein to preferentially bind to or modulate the activity of a CARK target molecule.

The cell-free assays of the present invention are amenable to use of both soluble and/or membrane-bound forms of isolated proteins (*e.g.*, CARK proteins or biologically active portions thereof). In the case of cell-free assays in which a membrane-bound form of an isolated protein is used it may be desirable to utilize a solubilizing agent such that the membrane-bound form of the isolated protein is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton[®] X-100, Triton[®] X-114, Thesit[®], Isotridecypoly(ethylene glycol ether)_n, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either CARK or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a CARK protein, or interaction of a CARK protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain

that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/ CARK fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test
 5 compound and either the non-adsorbed target protein or CARK protein, and the mixture incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above.

10 Alternatively, the complexes can be dissociated from the matrix, and the level of CARK binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a CARK protein or a CARK target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated
 15 CARK protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with CARK protein or target molecules but which do not interfere with binding of the CARK protein to its target molecule can be
 20 derivatized to the wells of the plate, and unbound target or CARK protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the CARK protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the
 25 CARK protein or target molecule.

In another embodiment, modulators of CARK expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of CARK mRNA or protein in the cell is determined. The level of expression of CARK mRNA or protein in the presence of the candidate compound is compared to the level of expression of CARK
 30 mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of CARK expression based on this comparison. For example, when expression of CARK mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of CARK mRNA or protein expression.
 35 Alternatively, when expression of CARK mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of CARK mRNA or protein expression. In one embodiment, CARK mRNA or protein expression is monitored in a cell based model for a cardiovascular

disorder, *e.g.*, a model of cardiac myocyte hypertrophy (Kariya, K *et al.* (1994) *J. Biol. Chem.*, 269:3775-3782). The level of CARK mRNA or protein expression in the cells can be determined by methods described herein for detecting CARK mRNA or protein.

In yet another aspect of the invention, the CARK proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with CARK ("CARK-binding proteins" or "CARK-bp") and are involved in CARK activity.

Such CARK-binding proteins are also likely to be involved in the propagation of signals by the CARK proteins or CARK targets as, for example, downstream elements of a CARK-mediated signaling pathway. Alternatively, such CARK-binding proteins are likely to be CARK inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a CARK protein is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a CARK-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the CARK protein.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a CARK protein can be confirmed *in vivo*, *e.g.*, in an animal, such as an animal model for a cardiovascular disorder, *e.g.*, a model of cardiac myocyte hypertrophy.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a CARK modulating agent, an antisense CARK nucleic acid molecule, a CARK-specific antibody, or a CARK-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an

agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

5

B. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

15

1. Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the CARK nucleotide sequences, described herein, can be used to map the location of the CARK genes on a chromosome. The mapping of the CARK sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease. The CARK gene has been mapped to human chromosome 1, within the atrioventricular canal defects (AVCD) locus (Sheffield, VC *et al.* (1997) *Human Molecular Genetics*, 6:117-121).

Briefly, CARK genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the CARK nucleotide sequences. Computer analysis of the CARK sequences can be used to predict primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the CARK sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (*e.g.*, human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but human cells can, the one human chromosome that contains the gene encoding the needed enzyme, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy

35

mapping of individual genes to specific human chromosomes. (D'Eustachio P. *et al.* (1983) *Science* 220:919-924). Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

- 5 PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the CARK nucleotide sequences to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a
- 10 CARK sequence to its chromosome include *in situ* hybridization (described in Fan, Y. *et al.* (1990) *Proc. Natl. Acad. Sci. USA*, 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries.

- Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one
- 15 step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical such as colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as
- 20 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma *et al.*, Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York 1988).

- 25 Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross
- 30 hybridizations during chromosomal mapping.

- Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a
- 35 gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland, J. *et al.* (1987) *Nature*, 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the CARK gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

2. Tissue Typing

The CARK sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the CARK nucleotide sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The CARK nucleotide sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1 or 7 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000

primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO:3 or 9 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from CARK nucleotide sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

3. Use of Partial CARK Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, *e.g.*, hair or skin, or body fluids, *e.g.*, blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, *e.g.*, PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.* another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1 or 7 are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the CARK nucleotide sequences or portions thereof, *e.g.*, fragments derived from the noncoding regions of SEQ ID NO:1 or 7 having a length of at least 20 bases, preferably at least 30 bases.

The CARK nucleotide sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, *e.g.*, brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such CARK probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, *e.g.*, CARK primers or probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

C. Predictive Medicine:

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining CARK protein and/or nucleic acid expression as well as CARK activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant or unwanted CARK expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with CARK nucleic acid and/or protein expression or activity. For example, mutations in a CARK gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with CARK nucleic acid and/or protein expression or activity. In one embodiment, CARK nucleic acid and/or protein expression or activity can be used in diagnostic and prognostic assays to detect and/or treat a cardiovascular disorder, *e.g.*, cardiac hypertrophy, congestive heart failure, a congenital heart defect, *e.g.*, an atrioventricular canal defect.

Another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of CARK in clinical trials.

These and other agents are described in further detail in the following sections.

1. Diagnostic Assays

An exemplary method for detecting the presence or absence of CARK protein or nucleic acid in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting CARK protein or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes CARK protein such that the presence of CARK protein or nucleic acid is detected in the biological sample. A preferred agent for detecting CARK mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to CARK mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length CARK nucleic acid, such as the nucleic acid of SEQ ID NO:1, 3, 7, or 9, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to CARK mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting CARK protein is an antibody capable of binding to CARK protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect CARK mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of CARK mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of CARK protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of CARK genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of CARK protein include introducing into a subject a labeled anti-CARK antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a serum sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting CARK protein, mRNA, or genomic DNA, such that the presence of CARK protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of CARK protein, mRNA or genomic DNA in the control sample with the presence of CARK protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of CARK in a biological sample. For example, the kit can comprise a labeled compound or agent capable of detecting CARK protein or mRNA in a biological sample; means for determining the amount of CARK in the sample; and means for comparing the amount of CARK in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect CARK protein or nucleic acid.

2. Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant or unwanted CARK expression or activity. As used herein, the term "aberrant" includes a CARK expression or activity which deviates from the wild type CARK expression or activity. Aberrant expression or activity includes increased or decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. For example, aberrant CARK expression or activity is intended to include the cases in which a mutation in the CARK gene causes the CARK gene to be under-expressed or over-expressed and situations in which such mutations result in a non-functional CARK protein or a protein which does not function in a wild-type fashion, *e.g.*, a protein which does not interact with a CARK ligand or one which interacts with a non-CARK ligand. As used herein, the term "unwanted" includes an unwanted phenomenon involved in a biological response such as proliferation or differentiation. For example, the term unwanted includes a CARK expression or activity which is undesirable in a subject.

The assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with a misregulation in CARK protein activity or nucleic acid expression, such as a cardiovascular disorder. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disorder associated with a misregulation in CARK protein activity or nucleic acid expression, such as a cardiovascular disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant or unwanted CARK expression or activity in which a test sample is obtained from a subject and CARK protein or nucleic acid (*e.g.*, mRNA or genomic DNA) is detected, wherein the presence of CARK protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant or unwanted CARK expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant or unwanted CARK expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a cardiovascular disorder. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder

associated with aberrant or unwanted CARK expression or activity in which a test sample is obtained and CARK protein or nucleic acid expression or activity is detected (*e.g.*, wherein the abundance of CARK protein or nucleic acid expression or activity is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant or
5 unwanted CARK expression or activity).

The methods of the invention can also be used to detect genetic alterations in a CARK gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by misregulation in CARK protein activity or nucleic acid expression, such as a cardiovascular disorder. In preferred embodiments, the methods include detecting, in a
10 sample of cells from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a CARK-protein, or the mis-expression of the CARK gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a CARK gene; 2) an addition of one or more nucleotides to a CARK gene;
15 3) a substitution of one or more nucleotides of a CARK gene, 4) a chromosomal rearrangement of a CARK gene; 5) an alteration in the level of a messenger RNA transcript of a CARK gene, 6) aberrant modification of a CARK gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a CARK gene, 8) a non-wild type level of a CARK-protein, 9)
20 allelic loss of a CARK gene, and 10) inappropriate post-translational modification of a CARK-protein. As described herein, there are a large number of assays known in the art which can be used for detecting alterations in a CARK gene. A preferred biological sample is a tissue or serum sample isolated by conventional means from a subject.

In certain embodiments, detection of the alteration involves the use of a probe/primer
25 in a polymerase chain reaction (PCR) (see, *e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, *e.g.*, Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the CARK-gene (see Abravaya *et al.* (1995) *Nucleic Acids Res.* 23:675-
30 682). This method can include the steps of collecting a sample of cells from a subject, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a CARK gene under conditions such that hybridization and amplification of the CARK-gene (if present) occurs, and detecting the presence or absence of an amplification product, or
35 detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, D.Y. *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. *et al.* (1988) *Bio-Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a CARK gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in CARK can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin, M.T. *et al.* (1996) *Human Mutation* 7: 244-255; Kozal, M.J. *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in CARK can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, M.T. *et al. supra*. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the CARK gene and detect mutations by comparing the sequence of the sample CARK with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, *e.g.*, PCT International

Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in the CARK gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or
 5 RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type CARK sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded
 10 duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the
 15 mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al.* (1988) *Proc. Natl Acad Sci USA* 85:4397; Saleeba *et al.* (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more
 20 proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in CARK cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an
 25 exemplary embodiment, a probe based on a CARK sequence, *e.g.*, a wild-type CARK sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

30 In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in CARK genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA*: 86:2766, see also Cotton (1993) *Mutat. Res.* 285:125-144; and Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-
 35 stranded DNA fragments of sample and control CARK nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected

with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen
5 *et al.* (1991) *Trends Genet* 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature,
10 for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753).

Examples of other techniques for detecting point mutations include, but are not
15 limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al.* (1989) *Proc. Natl Acad. Sci USA* 86:6230). Such allele specific
20 oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides
25 used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in
30 the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by
35 looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose

patients exhibiting symptoms or family history of a disease or illness involving a CARK gene.

Furthermore, any cell type or tissue in which CARK is expressed may be utilized in the prognostic assays described herein.

5

3. Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (*e.g.*, drugs) on the expression or activity of a CARK protein (*e.g.*, the modulation signaling pathways associated with cellular growth and differentiation) can be applied not only in basic drug screening, but also in clinical trials.

10 For example, the effectiveness of an agent determined by a screening assay as described herein to increase CARK gene expression, protein levels, or upregulate CARK activity, can be monitored in clinical trials of subjects exhibiting decreased CARK gene expression, protein levels, or downregulated CARK activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease CARK gene expression, protein levels, or
15 downregulate CARK activity, can be monitored in clinical trials of subjects exhibiting increased CARK gene expression, protein levels, or upregulated CARK activity. In such clinical trials, the expression or activity of a CARK gene, and preferably, other genes that have been implicated in, for example, a CARK-associated disorder can be used as a "read out" or markers of the phenotype of a particular cell.

20 For example, and not by way of limitation, genes, including CARK, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) which modulates CARK activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on CARK-associated disorders (*e.g.*, disorders characterized by deregulated cellular growth or differentiation), for example, in a
25 clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of CARK and other genes implicated in the CARK-associated disorder, respectively. The levels of gene expression (*e.g.*, a gene expression pattern) can be quantified by northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by
30 measuring the levels of activity of CARK or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during treatment of the individual with the agent.

35 In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting

the level of expression of a CARK protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the CARK protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the CARK protein, mRNA, or genomic DNA in the pre-administration sample with the CARK protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of CARK to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of CARK to lower levels than detected, *i.e.* to decrease the effectiveness of the agent. According to such an embodiment, CARK expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

D. Methods of Treatment:

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant or unwanted CARK expression or activity. With regards to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics", as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers the study of how a patient's genes determine his or her response to a drug (*e.g.*, a patient's "drug response phenotype", or "drug response genotype".) Thus, another aspect of the invention provides methods for tailoring an individual's prophylactic or therapeutic treatment with either the CARK molecules of the present invention or CARK modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant or unwanted CARK expression or activity, by administering to the subject a CARK or an agent which modulates CARK expression or at least one CARK activity. Subjects at risk for a disease which is caused or contributed to by aberrant or unwanted CARK expression or activity can be identified by, for example, any or

a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the CARK aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of CARK aberrancy, for example, a CARK, CARK agonist or CARK antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

2. Therapeutic Methods

Another aspect of the invention pertains to methods of modulating CARK expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell with a CARK or agent that modulates one or more of the activities of CARK protein activity associated with the cell. An agent that modulates CARK protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring target molecule of a CARK protein (*e.g.*, a CARK substrate), a CARK antibody, a CARK agonist or antagonist, a peptidomimetic of a CARK agonist or antagonist, or other small molecule. In one embodiment, the agent stimulates one or more CARK activities. Examples of such stimulatory agents include active CARK protein and a nucleic acid molecule encoding CARK that has been introduced into the cell. In another embodiment, the agent inhibits one or more CARK activities. Examples of such inhibitory agents include antisense CARK nucleic acid molecules, anti-CARK antibodies, and CARK inhibitors. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant or unwanted expression or activity of a CARK protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) CARK expression or activity. In another embodiment, the method involves administering a CARK protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted CARK expression or activity.

For example, stimulation of CARK activity is desirable in situations in which CARK is abnormally downregulated and/or in which increased CARK activity is likely to have a beneficial effect. Likewise, inhibition of CARK activity is desirable in situations in which CARK is abnormally upregulated and/or in which decreased CARK activity is likely to have a beneficial effect.

3. Pharmacogenomics

The CARK molecules of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on CARK activity (*e.g.*, CARK gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) CARK-associated disorders (*e.g.*, cardiovascular disorders) associated with aberrant or unwanted CARK activity. In conjunction with such treatment, pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a CARK molecule or CARK modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a CARK molecule or CARK modulator.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. *et al.* (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11):983-985 and Linder, M.W. *et al.* (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (*e.g.*, a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants.) Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high resolution map can be generated from a combination of some ten-million known single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast

majority may not be disease-associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach", can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drug's target is known (*e.g.*, a CARK protein of the present invention), all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Alternatively, a method termed the "gene expression profiling", can be utilized to identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (*e.g.*, a CARK molecule or CARK modulator of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a CARK molecule or CARK modulator, such as a modulator identified by one of the exemplary screening assays described herein.

4. Use of CARK Molecules as Surrogate Markers

The CARK molecules of the invention are also useful as markers of disorders or disease states, as markers for precursors of disease states, as markers for predisposition of disease states, as markers of drug activity, or as markers of the pharmacogenomic profile of a subject. Using the methods described herein, the presence, absence and/or quantity of the CARK molecules of the invention may be detected, and may be correlated with one or more biological states *in vivo*. For example, the CARK molecules of the invention may serve as surrogate markers for one or more disorders or disease states or for conditions leading up to disease states.

As used herein, a “surrogate marker” is an objective biochemical marker which correlates with the absence or presence of a disease or disorder, or with the progression of a disease or disorder (*e.g.*, with the presence or absence of a tumor). The presence or quantity of such markers is independent of the causation of the disease. Therefore, these markers may serve to indicate whether a particular course of treatment is effective in lessening a disease state or disorder. Surrogate markers are of particular use when the presence or extent of a disease state or disorder is difficult to assess through standard methodologies (*e.g.*, early stage tumors), or when an assessment of disease progression is desired before a potentially dangerous clinical endpoint is reached (*e.g.*, an assessment of cardiovascular disease may be made using cholesterol levels as a surrogate marker, and an analysis of HIV infection may be made using HIV RNA levels as a surrogate marker, well in advance of the undesirable clinical outcomes of myocardial infarction or fully-developed AIDS). Examples of the use of surrogate markers in the art include: Koomen *et al.* (2000) *J. Mass. Spectrom.* 35:258-264; and James (1994) *AIDS Treatment News Archive* 209.

The CARK molecules of the invention are also useful as pharmacodynamic markers. As used herein, a “pharmacodynamic marker” is an objective biochemical marker which correlates specifically with drug effects. The presence or quantity of a pharmacodynamic marker is not related to the disease state or disorder for which the drug is being administered; therefore, the presence or quantity of the marker is indicative of the presence or activity of the drug in a subject. For example, a pharmacodynamic marker may be indicative of the concentration of the drug in a biological tissue, in that the marker is either expressed or transcribed or not expressed or transcribed in that tissue in relationship to the level of the drug. In this fashion, the distribution or uptake of the drug may be monitored by the pharmacodynamic marker. Similarly, the presence or quantity of the pharmacodynamic marker may be related to the presence or quantity of the metabolic product of a drug, such that the presence or quantity of the marker is indicative of the relative breakdown rate of the drug *in vivo*. Pharmacodynamic markers are of particular use in increasing the sensitivity of detection of drug effects, particularly when the drug is administered in low doses. Since

even a small amount of a drug may be sufficient to activate multiple rounds of marker (*e.g.*, a CARK marker) transcription or expression, the amplified marker may be in a quantity which is more readily detectable than the drug itself. Also, the marker may be more easily detected due to the nature of the marker itself; for example, using the methods described
 5 herein, anti-CARK antibodies may be employed in an immune-based detection system for a CARK protein marker, or CARK-specific radiolabeled probes may be used to detect a CARK mRNA marker. Furthermore, the use of a pharmacodynamic marker may offer mechanism-based prediction of risk due to drug treatment beyond the range of possible direct observations. Examples of the use of pharmacodynamic markers in the art include:
 10 Matsuda *et al.* US 6,033,862; Hattis *et al.* (1991) *Env. Health Perspect.* 90:229-238; Schentag (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3:S21-S24; and Nicolau (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3:S16-S20.

The CARK molecules of the invention are also useful as pharmacogenomic markers. As used herein, a "pharmacogenomic marker" is an objective biochemical marker which
 15 correlates with a specific clinical drug response or susceptibility in a subject (see, *e.g.*, McLeod *et al.* (1999) *Eur. J. Cancer* 35(12):1650-1652). The presence or quantity of the pharmacogenomic marker is related to the predicted response of the subject to a specific drug or class of drugs prior to administration of the drug. By assessing the presence or quantity of one or more pharmacogenomic markers in a subject, a drug therapy which is
 20 most appropriate for the subject, or which is predicted to have a greater degree of success, may be selected. For example, based on the presence or quantity of RNA, or protein (*e.g.*, CARK protein or RNA) for specific tumor markers in a subject, a drug or course of treatment may be selected that is optimized for the treatment of the specific tumor likely to be present in the subject. Similarly, the presence or absence of a specific sequence mutation
 25 in CARK DNA may correlate CARK drug response. The use of pharmacogenomic markers therefore permits the application of the most appropriate treatment for each subject without having to administer the therapy.

This invention is further illustrated by the following examples which should not be
 30 construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are incorporated herein by reference.

EXAMPLES

EXAMPLE 1: IDENTIFICATION AND CHARACTERIZATION OF HUMAN AND RAT CARK cDNAs

In this example, the identification and characterization of the genes encoding human CARK (clone fchrf013f03) and rat CARK is described.

Isolation of the human and rat CARK cDNAs

The invention is based, at least in part, on the discovery of the human and rat genes encoding CARK. The human CARK gene was isolated from a cDNA library which was prepared from tissue obtained from subjects suffering from congestive heart failure of ischemic and idiopathic origin. Briefly, a cardiac tissue sample was obtained from a biopsy of four patients suffering from congestive heart failure. mRNA was isolated from the cardiac tissue and a cDNA library was prepared therefrom using art known methods (described in, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989). Positive clones were isolated from these libraries using appropriate primers.

The sequence of the positive clone was determined and found to contain an open reading frame. The nucleotide sequence encoding the human CARK protein comprises about 3025 nucleic acids, and has the nucleic acid sequence shown in Figure 1 and set forth as SEQ ID NO:1. The protein encoded by this nucleic acid comprises about 835 amino acids, and has the amino acid sequence shown in Figure 1 and set forth as SEQ ID NO:2.

A clone containing the rat CARK cDNA was also identified. The nucleotide sequence encoding the rat CARK protein comprises about 3026 nucleic acids, and has the nucleic acid sequence shown in Figure 5 and set forth as SEQ ID NO:7. The protein encoded by this nucleic acid comprises about 835 amino acids, and has the amino acid sequence shown in Figure 5 and set forth as SEQ ID NO:8.

Analysis of human CARK

A BLASTP 1.4 search (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the protein sequence of human CARK revealed that CARK is similar to the *C. elegans* protein containing similarity to ankyrin repeats and protein kinase motifs (Accession No. AF024491). This protein is approximately 42% identical (over CARK amino acids 126-370), 60% identical (over CARK amino acids 421-548), 65% identical (over CARK amino acids 582-686), 32% identical (over CARK amino acids 60-222), 75% identical (over CARK amino acids 381-404), 36% identical (over CARK amino acids 122-167), 48% identical (over CARK amino acids 10-36), 38% identical (over CARK amino acids 91-126),

52% identical (over CARK amino acids 556-678), 24% identical (over CARK amino acids 315-368), 25% identical (over CARK amino acids 74-161), and 29% identical (over CARK amino acids 219-259) at the amino acid level.

A BLASTN 1.4.9 search (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the
5 nucleotide sequence of human CARK revealed that CARK is similar to Homo sapiens
cDNA clone 1320937 (Accession No. AA758546). The CARK nucleic acid molecule is
100% identical to Homo sapiens cDNA clone 1320937 (Accession No. AA758546) over
nucleotides 2544 to 3010.

Analysis of primary and secondary protein structures, as shown in Figure 2, was
10 performed as follows: alpha, beta turn and coil regions, Garnier-Robson algorithm (Garnier
et al. (1978) *J Mol Biol* 120:97); alpha, beta, and turn regions, Chou-Fasman algorithm
(Chou and Fasman (1978) *Adv in Enzymol Mol* 47:45-148); hydrophilicity and
hydrophobicity plots, Kyte-Doolittle algorithm (Kyte and Doolittle (1982) *J Mol Biol*
15 157:105-132); alpha amphipathic and beta amphipathic regions, Eisenberg algorithm
(Eisenberg *et al.* (1982) *Nature* 299:371-374); flexible regions, Karplus-Schulz algorithm
(Karplus and Schulz (1985) *Naturwissens-Chafen* 72:212-213); antigenic index, Jameson-
Wolf algorithm (Jameson and Wolf (1988) *CABIOS* 4:121-136); surface probability plot,
Emini algorithm (Emini *et al.* (1985) *J Virol* 55:836-839).

The CARK protein was aligned with the A. Thaliana kinase 2 (Accession
20 No.Z97337), A. Thaliana kinase 3 (Accession No. AC003113), Arabidopsis thaliana
(Accession No. AL031135), C. elegans kinase (Accession No. AF024491), D. discoideum
protein tyrosine kinase (Accession No. A35670), D. discoideum protein tyrosine kinase
(Accession No. U01064), H. sapiens serine/threonine kinase (Accession No. Z48615),
human raf1 (Accession No. W13107), human Raf1 kinase (Accession No. R98215), and
25 soybean kinase (Accession No. M67449) using the Clustal method with a PAM250 residue
weight table. This alignment is shown in Figure 3.

Searches of the amino acid sequence of human CARK resulted in the identification
of several possible phosphorylation sites within the amino acid sequence of the CARK
polypeptide (SEQ ID NO:2). Protein kinase C phosphorylation sites were identified at
30 residues 208-210, 643-645, and 824-826; a cAMP and cGMP dependent protein kinase
phosphorylation site was identified at residues 18-21; casein kinase II phosphorylation sites
were identified at residues 11-14, 123-126, 224-227, 257-260, 293-296, 376-379, 461-464,
499-502, 615-618, 744-747, 765-768, 794-797, 805-808 and 829-832; and a tyrosine kinase
phosphorylation site was identified at residues 764-771 of human CARK. The search also
35 identified the presence of N-glycosylation site motifs at amino acid residues 194-197, 579-
582, 620-623, 735-738 and 827-830; N-myristoylation site motifs at amino acid residues 2-
7, 50-55, 78-83, 91-96, 133-138, 157-162, 190-195, 343-348, 363-368, 420-425, 526-531,

545-550 and 786-791; and a protein kinase ATP binding-region signature (Prosite Accession PS00107) at amino acid residues 469-477 of human CARK.

Analysis of rat CARK

5 A BLASTP 1.4 search (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the protein sequence of rat CARK revealed that CARK is similar to the *Homo sapiens* putative protein tyrosine kinase (Accession No. AF116826). Rat CARK is approximately 91% identical to the *Homo sapiens* putative protein tyrosine kinase (Accession No. AF116826) over amino acid residues 1-835. Rat CARK is also similar to the *C. elegans* protein containing
10 similarity to ankyrin repeats and protein kinase motifs (Accession No. AF024491). Rat CARK is approximately 41% identical to the *C. elegans* protein containing similarity to ankyrin repeats and protein kinase motifs over amino acid residues 39-720, and 39% identical over amino acid residues 10-370.

15 A BLASTN 1.4.9 search (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the nucleotide sequence of rat CARK revealed that CARK is similar to the *Homo sapiens* clone HH498 putative protein-tyrosine kinase mRNA (Accession No. AF116826). The rat CARK nucleic acid molecule is 81% identical to the *Homo sapiens* clone HH498 putative protein-tyrosine kinase mRNA (Accession No. AF116826) over nucleotides 27 to 2987, and 63% identical over nucleotides 2752 to 2999.

20 The rat CARK nucleic acid sequence was globally aligned with the human CARK nucleic acid sequence using the GAP program in the GCG software package, using a nwsgapdna matrix a gap weight of 12 and a length weight of 4. The results showed a 82.2% identity between the two sequences (see Figure 7).

25 The rat CARK protein sequence was globally aligned with the human CARK protein sequence using the GAP program in the GCG software package, using a Blosum 62 matrix and a gap weight of 12 and a length weight of 4. The results showed a 91.4% identity between the two sequences (see Figure 8).

Tissue Distribution of CARK mRNA

30 This Example describes the tissue distribution of CARK mRNA, as determined by Northern blot hybridization.

Northern blot hybridization with the RNA sample was performed under standard conditions and washed under stringent conditions, *i.e.*, 0.2 X SSC at 65°C. A DNA probe corresponding to human CARK (clone fchr013f03) was used. The DNA was radioactively
35 labeled with ³²P-dCTP using the Prime-It kit (Stratagene, La Jolla, CA) according to the instructions of the supplier. Filters containing human mRNA, human muscle mRNA, and rat mRNA (MultiTissue Northern blots #7760-1, #7765-1, and #7764-1, respectively, from

Clontech, Palo Alto, CA) were probed in ExpressHyb hybridization solution (Clontech) and washed at high stringency conditions according to the manufacturer's recommendations.

On the human mRNA blot which contained mRNA from heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas, CARK message was strongly detected in heart and faintly in skeletal muscle. On the human muscle mRNA blot which contained mRNA from skeletal muscle, uterus, colon, small intestine, bladder, heart, stomach, and prostate, CARK message was strongly detected in the heart. On the rat mRNA blot which contained mRNA from heart, brain, spleen, lung, liver, skeletal muscle, kidney and testis, CARK message was strongly detected in the heart.

Furthermore, on a cell line blot which contained mRNA from HeLa, 10T1/2, C2C12 myoblasts, C2C12 myotubes, rat neonatal cardiac fibroblasts, rat adult cardiac fibroblasts, rat neonatal cardiac myocytes and rat adult cardiac myocytes, CARK message was strongly detected in adult cardiac myocytes.

Chromosome Mapping of the CARK Gene

The CARK gene was found to map to human chromosome 1, within the atrioventricular canal defects (AVCD) locus (Sheffield, VC *et al.* (1997) *Human Molecular Genetics*, 6:117-121). Therefore, the CARK gene is a candidate for a congenital heart defect susceptibility gene, *e.g.*, a septal defect, a endocardial cushion defect, a vessel defect, or a valvular defect.

EXAMPLE 2: REGULATION OF CARK EXPRESSION IN CARDIAC MYOCYTES

This Example describes the regulation of CARK expression in a cell culture model of cardiac hypertrophy, and in cardiac myocytes stimulated with growth factors.

Briefly, primary cultures of neonatal cardiac myocytes were incubated for up to 24 hours in serum free minimal essential medium (MEM) supplemented with transferrin (10 µg/ml), insulin (10 µg/ml), and bovine serum albumin (1 mg/ml) (Kariya, K *et al.* (1994) *J. Biol. Chem.*, 269:3775-3782). The α1-adrenergic agonist L-phenylephrine (20 µM) was added during the incubation period to induce cardiac myocyte hypertrophy. Alternatively, cells were treated with angiotensin II (AII, 10 µM), insulin-like growth factor (IGF, 1 ng/ml), transforming growth factor-β (TGF-β, 1 ng/ml), or tumor necrosis factor-α (TNF-α, 1 ng/ml). RNA was harvested for Northern blot analysis.

Stimulation of cardiac myocytes with the hypertrophic agonist phenylephrine resulted in the down-regulation of CARK mRNA in a time dependent fashion, with CARK mRNA reaching almost undetectable levels between 12 and 24 hours following treatment. The expression of atrial natriuretic factor (ANF) and sarcoplasmic reticulum-Ca(2+)-ATPase

(SERCA2) which are known to be regulated during myocyte hypertrophy was also assessed. ANF mRNA was induced and SERCA2 mRNA was down-regulated following phenylephrine treatment. The gene encoding the CHO-B ribosomal protein was used as a control gene which is not regulated during hypertrophy.

5 The effect of phenylephrine on CARK regulation is specific, as treatment with AII, IGF, TGF- β and TNF- α did not have any effect on CARK expression. These growth factors also did not induce a state of hypertrophy, as they did not induce ANF expression.

10 The CARK gene is specifically downregulated in response to the α 1-adrenergic agonist L-phenylephrine in a cell culture model of cardiac hypertrophy, and may thus play a role in cellular signal transduction pathways that regulate cardiac myocyte growth and differentiation. The regulation of the CARK gene during cardiac myocyte hypertrophy may be useful for identifying gene regulatory elements (*e.g.*, elements within the regulatory, non-coding sequences of the CARK gene) that modulate cardiac gene transcription during hypertrophic growth.

15 In another study, CARK gene expression was assessed by RT-PCR in cardiac myocytes treated with serum, TGF- β , and BMP-4 for 5 minutes, 2 hours, and 6 hours. Stimulation of cells with TGF- β had little effect on CARK gene expression, whereas treatment with BMP-4 and 10% fetal bovine serum resulted in a decrease in CARK expression.

20

 In another study, CARK expression in tissue samples from humans subjects with congestive heart failure was analyzed by TaqMan® Quantitative Polymerase Chain Reaction.

25 Briefly, probes were designed by PrimerExpress software (PE Biosystems) based on the sequence of the CARK gene. Each CARK gene probe was labeled using FAM (6-carboxyfluorescein), and the β 2-microglobulin reference probe was labeled with a different fluorescent dye, VIC. The differential labeling of the target gene and internal reference gene, thus, enabled measurement in the same well. Forward and reverse primers and probes for both the β 2-microglobulin and the target gene were added to the TaqMan® Universal
30 PCR Master Mix (PE Applied Biosystems). Although the final concentration of primer and probe could vary, each was internally consistent within a given experiment. A typical experiment contained 200nM of forward and reverse primers plus 100nM of probe for β -2 microglobulin and 600 nM of forward and reverse primers plus 200 nM of probe for the target gene. TaqMan matrix experiments were carried out using an ABI PRISM 7700
35 Sequence Detection System (PE Applied Biosystems). The thermal cycler conditions were as follows: hold for 2 minutes at 50°C and 10 minutes at 95°C, followed by two-step PCR for 40 cycles of 95°C for 15 seconds followed by 60°C for 1 minute.

A comparative Ct method was used for the relative quantitation of gene expression. The following method was used to quantitatively calculate CARK gene expression in the various samples relative to β -2 microglobulin expression in the same sample. The threshold cycle (Ct) value was defined as the cycle at which a statistically significant increase in fluorescence is detected. A lower Ct value was indicative of a higher mRNA concentration. The Ct value of the CARK gene was normalized by subtracting the Ct value of the β -2 microglobulin gene to obtain a Δ Ct value using the following formula:

$$\Delta Ct = Ct_{\text{CARK}} - Ct_{\beta\text{-2 microglobulin}}$$

Expression was then calibrated against a cDNA control sample containing no template. The Δ Ct value for the calibrator sample was then subtracted from Δ Ct for each tissue sample according to the following formula:

$$\Delta\Delta Ct = \Delta Ct_{\text{sample}} - \Delta Ct_{\text{calibrator}}$$

Relative expression was then calculated using the arithmetic formula given by $2^{-\Delta\Delta Ct}$.

Similarly, TaqMan® analysis was used to assess human atrial natriuretic factor (ANF) expression.

CARK gene expression was increased in ischemic heart tissue samples, including samples derived from the left ventricle, as compared to normal tissue controls (see Figure 9). ANF gene expression was induced in familial hypertrophic cardiomyopathy samples from the left atrium. These data indicate that CARK may play a role in the regulation of cardiac cell growth and/or differentiation and the pathogenesis of cardiovascular disorders, *e.g.*, congestive heart failure and cardiac hypertrophy.

EXAMPLE 3: DUAL KINASE ACTIVITY OF CARK POLYPEPTIDES

This example, describes the dual kinase activity of CARK polypeptides *in vitro* and *in vivo*.

GST-CARK fusion proteins were prepared for both the full length human CARK gene and a polypeptide comprising the CARK kinase domain (KD) (amino acid residues 424 to 835 of SEQ ID NO:2). Briefly, Sf9 cells were grown in 6 well plates, transfected with GST-CARK constructs, and the GST-tagged proteins were harvested from the cells as follows. Cells were harvested and stored at -80°C until ready for use. Cells were thawed on ice, and 500 μl of lysis buffer (1% NP-40, 50 mM HEPES pH7.5, 5% glycerol, 150 mM NaCl, 1 mM EDTA, 1:100 protease inhibitor cocktail (Calbiochem #539131)) was added to each well. Cells were scraped from the plate, transferred to microfuge tubes, vortexed, and incubated on ice for 30 minutes. Subsequently, the cells were subjected to several rounds of freeze-thaw lysis. The samples were then spun at high speed for 10 minutes at 4°C in a microfuge. The supernatants were transferred to new tubes, and incubated with 75 μl of glutathione sepharose 4B (Amersham 17-0756-01) for 1 hour at 4°C . The resin was

pelleted, the supernatant removed, and the resin washed 3 times with lysis buffer. The samples were eluted three times by incubation with 50 µl reduced glutathione for 10 minutes at room temperature. In some instances the samples were dialyzed twice at 4°C in 1 litre of dialysis buffer containing 50 mM HEPES pH7.5, 10% glycerol, 100 mM NaCl, 1 mM EDTA, 20 mM beta-glycerophosphate, 0.1 mM Na₃VO₄, 1mM PMSF and 1 mM DTT.

The kinase activity of the purified GST-CARK and GST-CARK-KD proteins was tested in an *in vitro* kinase assay. Briefly, the GST-CARK fusion proteins were incubated in a reaction mixture containing kinase buffer (20 mM MOPS pH7, 20 mM MgCl₂, 5 mM NaF, 1 mM sodium orthovanadate, 5 mM DTT), and ATP, separated by SDS-PAGE, and subjected to Western blot analysis using an anti-phosphotyrosine antibody (4G10). As shown in Figure 10, both the GST-CARK and GST-CARK-KD polypeptides were capable of autophosphorylation on tyrosine residues, and this activity was ATP dependent. A tyrosine phosphorylation site is predicted at amino acid residue Y771 of the CARK polypeptide (amino acid residues 764-771 of SEQ ID NO:2 and 8).

In addition, the kinase activity of the GST-CARK and the GST-CARK-KD proteins towards heterologous substrates, *e.g.*, inactive Erk-2, GST-Elk-1, GST-c-jun, GST-ATF2, myelin basic protein, H1 histone and Phas-1 (4E-BP-1), was tested in an *in vitro* kinase assay. The kinase assay reactions conditions (25 µl total volume) were as follows:

- 2.5 µl substrate (5 µg)
- 5.0 µl ATP mix
 - (1.5 µl kinase buffer/2.5 µl cold ATP (10 µM stock)/1 µl γ³²P ATP (10 µCi/µl stock))
- 10 µl purified GST-CARK protein (in kinase buffer)
- 7.5 µl kinase buffer
- (20 mM MOPS pH7/20 mM MgCl₂/5 mM NaF/1 mM sodium orthovanadate/5 mM DTT)

Reactions were incubated at 30°C for 30 minutes, mixed with an equal volume of 2X Laemmli sample buffer, heated to 95°C for 5 minutes. 20 µl samples were run on a 4-20% gradient gel. The GST-CARK polypeptides phosphorylated the H1 histone, myelin basic protein, ATF-2 and Phas-1 substrates, thus demonstrating serine/threonine kinase activity. Moreover, CARK kinase activity was specifically inhibited in a dose dependent fashion by the broad range serine/threonine kinase inhibitors staurosporine and K252, whereas the inhibitors bisindolylmaleimide, H-89, PKG, ML-7, KN-93, and H-7, specific for PKC, PKA, PKG, MLCK, CaM kinase, and MLCK/PKA/PKC/PKG, respectively, had no effect.

In another study, 293 cells were transfected with myc-tagged full length CARK and dominant negative CARK containing a lysine to arginine substitution at amino acid residue 490 of SEQ ID NO:2 (K490R), and expression of the CARK polypeptides was confirmed by

western blot analysis using an anti-myc antibody. Cell lysates were immunoprecipitated with an anti-myc antibody, transferred to a nitrocellulose membrane, and blotted with an anti-phosphotyrosine antibody. The myc-CARK polypeptide was autophosphorylated on tyrosine residue(s) whereas dominant negative CARK was not autophosphorylated (Figure 11).

EXAMPLE 4: SUBCELLULAR LOCALIZATION OF RECOMBINANT CARK PROTEIN

Vero 2-2 cells were transfected with a full length GFP-CARK construct and subcellular localization of the GFP-CARK polypeptide was observed by fluorescence microscopy. The GFP-CARK polypeptide was expressed in the cytoplasm of the cell.

EXAMPLE 5: EXPRESSION OF RECOMBINANT CARK PROTEIN IN CARDIAC MYOCYTES

To express the CARK gene in cardiac myocytes, the human CARK cDNA was cloned into a Herpes Simplex Virus (HSV) vector. In addition to wild-type CARK, a dominant negative (DN) mutant CARK protein containing a lysine to arginine substitution at amino acid residue 490 of SEQ ID NO:2 (K490R) was expressed. Additional CARK mutants that were tested included hyperactive CARK proteins containing either a lysine to aspartic acid substitution (K490D) or a lysine to glutamic acid substitution (K490E) at amino acid residue 490 of SEQ ID NO:2; a CARK active site mutant containing glycine to alanine substitutions at amino acid residues 470, 472 and 475 of SEQ ID NO:2; and a CARK myristoylation mutant containing a deletion of amino acid residues 2 to 7 of SEQ ID NO:2 which comprises a predicted N-myristoylation site, and which may interfere with CARK subcellular localization (*e.g.*, membrane localization) and signaling activity.

Rat neonatal cardiac myocytes were infected with either empty HSV vector, HSV-GFP (green fluorescent protein), HSV-CARK, or HSV-CARK (DN). Infected cells were analyzed for CARK gene expression, and cellular phenotypes including intracellular calcium levels, protein synthesis, expression of hypertrophic markers (*e.g.*, atrial natriuretic factor (ANF)), and expression of certain other genes of interest, (*e.g.*, oncogenes, tumor suppressors, cell cycle regulators, transcription factors, cell surface antigens, or intracellular signal transduction modulators and effectors, or those genes involved in stress response, ion channels and transport, apoptosis, DNA repair, receptors, or cell-cell communications). ANF expression was induced in rat neonatal cardiac myocytes infected with HSV-CARK as compared to control cells as assessed by RT-PCR. As shown in Figure 12, the expression of dominant negative CARK inhibited *de novo* protein synthesis induced by phenylephrine (40

μM) or serum (10% fetal bovine serum) in neonatal cardiac myocytes as assessed by the incorporation of ¹⁴C-phenylalanine.

EXAMPLE 6: GENERATION OF CARK TRANSGENIC ANIMALS

5

A CARK cDNA was ligated to the cardiac-specific alpha-myosin heavy chain promoter using art known techniques. This construct was microinjected into the male pronucleus of fertilized mouse eggs, which were subsequently implanted into a pseudopregnant foster mother. Founder mice were bred with non-transgenic cohorts for the propagation of individual lines.

10

The presence of the CARK transgene in the transgenic mice was confirmed by Southern analysis of genomic DNA. The phenotype of the transgenic mice was also assessed indicating that the CARK transgenic animals have enlarged hearts. Specifically, an average heart-to-body ratio of 6.42 was observed for the CARK transgenic animals, as compared to an average heart-to-body ratio of 4.77 that was observed for the wild-type animals. Prominent cardiac myocyte hypertrophy was also observed in the CARK transgenic animals.

15

EXAMPLE 7: EXPRESSION OF RECOMBINANT CARK PROTEIN IN BACTERIAL CELLS

20

In this example, CARK is expressed as a recombinant glutathione-S-transferase (GST) fusion polypeptide in *E. coli* and the fusion polypeptide is isolated and characterized. Specifically, CARK is fused to GST and this fusion polypeptide is expressed in *E. coli*, e.g., strain PEB199. Expression of the GST-CARK fusion protein in PEB199 is induced with IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads. Using polyacrylamide gel electrophoretic analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined.

25

30

EXAMPLE 8: EXPRESSION OF RECOMBINANT CARK PROTEIN IN COS CELLS

35

To express the CARK gene in COS cells, the pcDNA/Amp vector by Invitrogen Corporation (San Diego, CA) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA fragment encoding the entire CARK protein and an HA tag (Wilson *et al.* (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the polylinker region of the vector, thereby placing the expression of the recombinant protein under the control of the CMV promoter.

To construct the plasmid, the CARK DNA sequence is amplified by PCR using two primers. The 5' primer contains the restriction site of interest followed by approximately twenty nucleotides of the CARK coding sequence starting from the initiation codon; the 3' end sequence contains complementary sequences to the other restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last 20 nucleotides of the CARK coding sequence. The PCR amplified fragment and the pCDNA/Amp vector are digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly, MA). Preferably the two restriction sites chosen are different so that the CARK gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5a, SURE, available from Stratagene Cloning Systems, La Jolla, CA, can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

COS cells are subsequently transfected with the CARK-pcDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. The expression of the CARK polypeptide is detected by radiolabelling (³⁵S-methionine or ³⁵S-cysteine available from NEN, Boston, MA, can be used) and immunoprecipitation (Harlow, E. and Lane, D. *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988) using an HA specific monoclonal antibody. Briefly, the cells are labelled for 8 hours with ³⁵S-methionine (or ³⁵S-cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

Alternatively, DNA containing the CARK coding sequence is cloned directly into the polylinker of the pCDNA/Amp vector using the appropriate restriction sites. The

resulting plasmid is transfected into COS cells in the manner described above, and the expression of the CARK polypeptide is detected by radiolabelling and immunoprecipitation using a CARK specific monoclonal antibody.

5 Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.